

23 July 1981

Cladistics and evolution on display

How should museums of natural history exhibit specimens? Is cladistics a proper framework? And is it too agnostic—soft on Neo-Darwinism? The last contribution to the correspondence on these topics is on page 403.



The long correspondence about the British Museum (Natural History) — whose name should be changed — has raised more issues than have been resolved, so that on the face of things this is an inappropriate time to bring an end to the affair. In reality, however, the time has come when there is very little fresh that can be said about the issues originally raised by L.B. Halstead (whose final contribution appears on page 403) and the other issues that have accreted over the months. (It is to be hoped that correspondents whose letters have not been published will appreciate that their arguments have mostly already been made by others.) The need now is that the contentious issues should not be forgotten — and that the museum should not think that the ending of the correspondence implies acquiescence in all aspects of its policy.

The role of the technique of classification called cladistics in the museum's exhibition policy evidently appears to many at the museum to be one of the chief issues to have been raised. The letter from Dr Roger Miles, the director of the Public Services Department at the museum (which plans and mounts exhibitions), and his colleague Dr G.C.S. Clarke on page 402 defends the decision, in two of the recent exhibitions, that the relationship between specimens displayed should be described in cladistic terms — what first set Halstead off. The decision is justified on "educational grounds", presumably the argument that visitors to these exhibitions should be provided both with an understanding of how phylogenetic relationships may be inferred by objective means (which is one of the undoubted virtues of cladistics, at least when meaningful correlations are obtainable) and with a sense of the uncertainty running through phylogeny. This is consistent with the principles on which curriculum development in high-school science have been constructed in the past twenty years. But museums are not high-school classrooms, in which good teachers can ensure that students appreciate the basis on which their work is carried out. Most visitors to museums see what is displayed but once. To fail to record, for such people, that in the opinion of most palaeoanthropologists, *Homo erectus* is the most likely precursor of *Homo sapiens*, is to lose an important opportunity — and to avoid answering questions in many visitors' minds about the importance of recent discoveries in East Africa. And to fail to make explicit the principles on which these exhibitions were designed is to ask for trouble from people such as Halstead, caught up as they are with the argument about the validity of cladistics as a way of classification and as a cloak for ideology.

Technically, cladistics has respectable origins, not least in the development of numerical taxonomy twenty years ago. Inevitably, however, the precision of a cladistic classification of fossil material is limited by the amount and variety of material available for analysis. Out-and-out cladists may hold that phylogenetic relationships not demonstrated by their techniques have no place in the interpretation of the fossil record and much of what they say makes philosophical sense. On present information, for example, there is no way of telling whether *Archaeopteryx* is the ancestor of all birds or (more probably) an early version of a bird. But pedantry can be carried too far, especially in public exhibitions. Although there may be cladistic doubt (represented by some such statement as " $P < 0.1$ ") about the relationship

between *Homo erectus* and *Homo sapiens*, do not the methods of classical palaeoanthropology, with all the appropriate qualifications, deserve a place in public exhibitions because they are stimulating, provoking?

A more serious difficulty, in the course the museum has elected to follow, is that the technique is best at spotting branching points in familial relationships, and says little about gradual transformations within a single population so profound that they are tantamount to a qualitative transformation. No wonder that Halstead fears that cladism is a refuge for those who believe in punctuated equilibrium. But, while he may be able to marshal substantial evidence for the moderate assertion that many Marxist evolutionists are cladists, the proposition that all cladists are Marxists — which has not seriously been suggested by anybody, but only inferred — has no place in logic. It would in any case, as Sir Andrew Huxley said in his attack on *Nature* on 27 May, be intolerable that such a reading of events should be held to be part of the case against cladistics.

The other serious objection to the modelling of public exhibitions within the framework of cladistics is that the framework is inherently agnostic. When the numerical correlations are too small to be significant when assessed against arbitrary criteria, relationships cannot be asserted. Although, even in exhibitions mounted for the general public, statements such as that "the case for thinking that *Homo erectus* was the precursor of *Homo sapiens* is not proved" may have a place, part of the trouble at the museum (to which there has been no reply) is that agnosticism has gone too far. Did evolution happen? The museum has cladistically not committed itself. Is Darwinism a "proper" scientific theory, in some sense of that word? The museum has withdrawn the obscure film that suggested to anybody who paused in passing that it is all a gigantic muddle; people will be looking, in the months ahead, at the museum's strategy for retreating from this gaffe. The difficulty, with which everybody must sympathize, is that all museums of natural history are used not merely by those who wish to answer questions in their minds about evolution and the like but also by those who believe that Darwinism is an abomination.

Several other issues remain to be resolved. The aesthetics of the museum's plans for the future have been questioned, and rightly. Is the new building really necessary? Is the push-button public exhibition the only acceptable modern style? Most certainly not. Was Sir Andrew Huxley's attack on *Nature* (for publishing alleged irrelevancies) just? Readers must judge for themselves. It is understandable that a trustee of the museum should defend members of the museum staff from supposedly (but imagined) personal attacks at the celebration of the museum's centenary. The issues raised are not, however, irrelevancies, even if not all of them are relevant to the museum. The defence would have been more persuasive if there had been less of it. And it was unfair to suggest that *Nature's* interest was by sensation to enlarge its circulation.

In all the argument of the past several months, perhaps the most important issue has been curiously neglected. Like many other invaluable institutions, the Natural History Museum is staffed by scientists who are public servants. This circumstance is not usually restrictive. Members of the museum's staff ordinarily function as



scholars without impediment. On the question of the public policy of the institution for which they work, however, civil service employees are less than free. In the past several months, many among the museum's staff have sympathized with many of the complaints made against their institution, but have been gagged. The problem is not novel. Scientists working in defence establishments may, for example, differ from official policy on defence procurement, but may be unable to say so publicly. Those working on the development of the Anglo-French supersonic aircraft Concorde may have considered at the time that the project would be a waste of money and resources, and would by now have been justified by events. In such circumstances, however, it is understandable (if reprehensible) that laboratories should require their employee-scientists to keep their silence. The case for insisting on such a policy at an institution such as a natural history museum, whose public reputation stems from the supposed academic independence and integrity of its members, is by comparison non-existent. It is to be hoped that the museum's trustees will now give this problem the attention it deserves, and urgently. For the correspondence of the past months shows that there is a great depth of seemly and legitimate professional interest in Darwinism, cladism and related matters. And it is not long until next April, when it will be time to mark the centenary of Charles Darwin's death.

Too much government

The British government has once again slipped into the trap of taking too detailed an interest in the management of technology, this time in the development of the British nuclear power programme. Last week, the Department of Energy published its reply (see page 398) to the Select Committee on Energy's report which, among other things, advocated that attention should be paid, even at this late stage, to the building of heavy water reactors on the Canadian pattern. In doing so, the committee was keeping faith with the traditions established by its predecessors (principally the now defunct Select Committee on Science and Technology) in hankering after reactor designs that nobody with direct responsibility for generating electricity wants to build. (Nearly a decade ago, the committee was pushing for the steam generating heavy water reactor, a complicated device using heavy water as a neutron moderator and pressurized ordinary water as a coolant, which was in 1977 discovered to be too complicated to fabricate.)

The Department of Energy rightly said last week that nothing has happened in the past four years to suggest that Canadian heavy water reactors would be economic in British circumstances, and thus rejected the committee's chief technical recommendation. As the department responsible for the Central Electricity Generating Board, the principal owner of reactors in the British electricity supply network, the department is well within its rights in making this opinion public. But it goes too far when it declares that the next thermal reactor ordered for the British network "should be the Pressurized Water Reactor". By this declarative statement of technical policy, the Department of Energy has committed the British government not merely to a technical policy that it (or its successor) may regret, but to a role in the management of technology that it should not be playing.

The circumstances are clear. The nationalized industry with a near-monopoly right to generate electricity south of the Scottish border is the Central Electricity Generating Board, much criticized by the Select Committee on Energy for its over-optimistic forecasts of electricity demand and for its imprecise calculations of the relative costs of generating electricity from different fuels. Constitutionally, the board has an unambiguous technical responsibility for deciding how best to generate electricity. If, with the passage of time, the board is found to be incompetent, either technically or in its economic judgements, the Department of Energy has a right, even a responsibility, to fire the board and appoint another. But if the department also sets itself up as the body that decides what kinds of reactors should be built,

its freedom to exercise its own constitutional responsibilities must be compromised. Worse, the government itself, and not the generating board, will become the target for the torrent of complaint that will be mobilized at next year's public inquiry at the proposal to "recreate Three Mile Island in Suffolk" (for where the generating station is destined). The result will be that the government's essential quasi-judicial role as the arbiter between the generating board and public protesters will also be undermined. How can it be otherwise when the department in its reply to the select committee uses forms of words such as "the government believes" that the next British reactor ordered should be a pressurized water reactor?

This complaint is not mere pendency. Forms of words matter, and in this case all too clearly reveal that the Department of Energy has not clarified for itself its relationship with the electricity-generating industry. The result is that the department's announcement of the form of the public inquiry about the first British reactor of this type has a hollow ring. Last week's document says that the inquiry is to be "wide ranging", that documents describing the studies that will have been made of the safety of the British reactor should be published well in advance, and that objectors should have time in which to prepare their case. The department rejects the select committee's proposal that a time limit should be fixed for the inquiry. All this is sensible enough if the objective is to reconvert British public opinion to nuclear power. The snag is that the department, having announced in advance its "belief" in the generating board's plan, will be accused at the inquiry of being a prejudiced judge in the case. This is a surprising error for a government which, at its election two years ago, advocated that government should be less obtrusive.

Over-civil servants

The trouble with the British government's response last week to the inter-departmental review of the scientific civil service published nearly a year ago (see *Nature* 287, 264; 1980) is that much of what it says has been said before. Yes, steps will be taken to help people recruited into the public service as scientists to move to administrative work. Yes, arrangements will be worked out to make the careers of those who stay as working scientists more interesting and challenging. Yes, steps will be taken to see whether government departments could make fuller use of scientists on short-term appointments. Yes, steps will be taken to define the roles of chief scientists in government departments (and then to ensure that these roles are carried out). Yes, steps will be taken to improve the literature handed out to potential recruits to the British scientific civil service (and departments will be asked to explain themselves to the Civil Service Department if they make a habit of recruiting highly qualified people to relatively lowly jobs). Yes, yes; the future will indeed be different from the past and better, more enlightened, more flexible, more challenging.

Brave words do not necessarily come true, as the negligible response of successive governments to the report of Lord Fulton's Royal Commission on the Civil Service more than a decade ago has shown. Briefly, however, the British civil service (and any other public service) has two overriding problems to solve in its relationship with scientists. First, how is the appalling lack of scientific understanding in the formulation of general policy to be made good? And how are those employed exclusively as scientists to be kept interested and alert? The simplest course would solve both problems. Let the recruitment of professional scientists (PhDs for example) to the public service be at the outset for an initial period of seven years. Let it be understood that those dispensed with at the end of this period are given a substantial gratuity, say a year's salary. And let it be understood that those kept on should either help to make policy or function as professional scientists, with the proviso that while they may be required confidentially to advise the departments that employ them on technical matters, they should be free to take part in the affairs of the intellectual professions that gave them skill (see above).

US industry moves into biotechnology

Du Pont sets up new lab, others to come

Washington

Reluctant to commit themselves to the commercial promise of the "new biology" only a few years ago, the United States chemical giants are now queuing up to jump in the deep end. Last week the largest of all, Du Pont, announced plans for a new \$85 million life sciences complex, designed for 700 scientists and technical personnel, to be built near the company's headquarters in Wilmington, Delaware.

One of Du Pont's closest rivals, Monsanto, has already established what the company describes as a "free-standing molecular biology center", which will have "a larger component of basic research than at most industrial companies". According to Monsanto chairman, Mr John Hanley, genetic engineering is likely to become a major branch of the company in the future; "we're up to the eyeballs in molecular biology", he says.

Both companies already have substantial contracts with outside research institutions. Last month, for example, Du Pont announced a five-year, \$6 million agreement with Harvard Medical School (see *Nature* 16 July, p. 191) to conduct research on molecular genetics, with the company receiving exclusive licences to market any resulting products. Monsanto signed a similar deal with the same institution, though with a slightly different focus, six years ago; recently it has been working with Genentech on bovine and porcine hormones, looking at ways of producing more meat with less feed.

At the same time, however, major chemical companies such as Du Pont and Monsanto are building up their in-house capacity at a rate which is already producing major recruitment difficulties.

Du Pont's life sciences complex is part of a plan to raise from 10 to 20 per cent the proportion of gross revenues which the company earns from sales of pharmaceutical, biomedical and agricultural products. According to the company's new chairman, Mr Edward Jefferson, it intends to spend 21 per cent of its \$570 million research budget next year on the life sciences, and that figure is expected to increase over the next five years.

The new complex will include a 250,000-square-foot facility for health sciences research and a 100,000-square-foot expansion of existing plant laboratories at the company's experimental station in Delaware. Completion is

scheduled for December 1983, and the company intends to expand the scope of current research not only in molecular genetics, but also in fields such as immunology, neurobiology, cardiology, gerontology, plant biology, plant pathology and entomology.

Du Pont's announcement seems to have been partly precipitated by rumours that the company's current attempts to take over the oil company Conoco could interfere with Du Pont's recent moves to diversify away from petroleum-based products. Mr Jefferson has denied such rumours, claiming that the company intends to continue its aggressive thrust into the life sciences for their own sake—and for the profit they may bring.

Several petroleum companies are themselves beginning to enter the fields. Atlantic Richfield has set up a genetic engineering laboratory which company officials describe as "our most academic laboratory"; novel arrangements being discussed include two-year research appointments, and lengthy sabbaticals for research workers who wish to spend time at universities.

Waiting in the wings is the Exxon Research and Engineering Company, the research wing of Exxon Corporation, with an annual research budget of \$489 million—almost half that of the National Science Foundation. Company officials declined

to comment last week on the company's plans, saying that no announcement was expected for several months; however, there are rumours that Exxon is working on an arrangement that it likes to compare with the position of the Bell Laboratories in telecommunications research.

The broad potential impact of life sciences research in medicine is revealed in a report, entitled *Forecast of Emerging Technologies*, soon to be released by the Food and Drug Administration, of a study in which 190 experts both inside and independent of the agency were asked to predict which new medical technologies were likely to emerge within the next fifteen years.

Almost a quarter of the 168 technologies listed involved basic areas of genetic engineering and computer technology. Eight technologies accounted for one-third of the citations, headed by hybridoma technology, nuclear magnetic resonance imaging and the use of recombinant DNA to produce interferon.

Meanwhile at least one university has decided to provide financial support to faculty members attempting to exploit the result of their research in molecular biology. The Michigan State University Foundation has announced that it is contributing \$100,000 to help set up a new company, called Neogen, which is expected to specialize in the genetic engineering of plants.

David Dickson

Research council plans academic rescue

Britain's Science and Engineering Research Council is to come to the aid of outstanding academics threatened with redundancy as a result of cuts in the universities' income. Last week, the council decided that it would help academics in departments marked for closure to transfer to other universities, and would pay their salaries for up to ten years if necessary. The council is calculating that the universities should then have reorganized themselves and that the University Grants Committee should be able to take up the bill again.

The council is particularly anxious that good individuals and groups should not be lost simply because they happen to work in those universities faced with the largest cuts in their income. It will be looking for academics that it considers especially valuable who are about to lose their jobs and may also consider setting up special transfer fellowships for which threatened academics will be eligible to apply.

As yet, the council has little idea of how many individuals it will want to support or how much money it should set aside. Estimates will have to wait at least until the autumn, according to Sir Geoffrey Allen, chairman of the council, when universities have a better idea of where economies will fall. In the meantime, the council has increased the number of replacement

fellowships, by which a senior academic returns to research leaving a tenured post for a younger person, from 15 to 20 per year. The council, which will pay the net cost for five years, expects to spend £500,000 a year on the scheme.

Sir Geoffrey's main concern is that the cuts in university income will further undermine the dual support system for research, whereby the University Grants Committee provides funds for staff and well equipped laboratories and the research councils award grants for specific projects. The gradual undermining of the universities' income over recent years has already weakened the research base, according to Sir Geoffrey, who is anxious that the restructuring of the university system caused by the cuts should be designed to strengthen the research base to near its former level rather than weaken it further.

He would like to see individual universities concentrate their efforts to a greater extent than implied by the University Grants Committee in its advice to them, preferring a smaller but stronger research base to a larger, weaker one. Although those views are shared by the University Grants Committee, which has attempted to confine the misery caused by the cuts by being selective, Sir Geoffrey believes that the selectivity could and should go further.

Judy Redfearn

French universities

Decision postponed

The 74 French universities, many of them created in the reform which followed the student rebellion of 1968, are in suspense. Does the new Socialist government love them or not?

The Minister of Education, M. Alain Savary, is keeping his own counsel. He has received representations from all sectors of the universities, from unions to presidents, since his invitation to the universities to appeal against the recently-published list of *habilitations* — a list of qualifications that each university is allowed to offer. The list was conceived largely by the previous Giscardian minister, Mme Alice Saunier-Seïté, and eliminated many old courses and refused the introduction of some new ones. A second, supplementary list was promised for last Saturday — but it did not materialize.

Now, says the ministry, it will appear at the beginning of August, when conveniently most of France is on holiday. And at the end of the vacation, M. Savary is to call unions and university presidents to a discussion, where he will attempt to reach agreement on new rules for creating university research and teaching programmes.

The implication may be that the new list will not be everything the universities want — which is approval for all the courses they propose. For defying Savary's call for self-restraint, the universities have thrown themselves solidly behind their original lists, so heavily pruned by Saunier-Seïté.

Nevertheless, they could have done little else, for Savary has given them no guidelines for what to keep and what to cut.

In the end, it may be the level of political pressure that has been brought on him that will determine which courses are approved. For example, at the Université de Haut-Bretagne (Rennes II), it appears that a course on the Breton language will be restored in the second list because of pressure from a group of Breton parliamentarians. After all, the government is pledged to respect and increase regional autonomy. And the new emphasis on the oceans (France has for the first time a minister of the sea) may save a course of oceanography at the Université de Toulon et du Var.

Savary certainly declared himself appalled at the state of university politics when he arrived in office, and it has not all been attributable to his predecessor. According to one senior scientific official, the university system had got out of hand after 1968, and needed to be trimmed. Ninety-five per cent of Saunier-Seïté's cuts were well judged, he thought.

However, such outspoken officials and heads of department may find themselves the symbolic sacrifices of the new government, to satisfy the multitude while the new politics is being thrashed out. At

the beginning of this month, one of the teaching unions (SGEN-CFDT) called for the resignation of certain heads of department who were still actively supporting Saunier-Seïté's point of view; and now it seems the Council of Ministers has a list of 27 names before it. These counter-revolutionaries may have to go, it is implied.

Robert Walgate

UK nuclear energy

Maintaining course

The British government remains convinced of the wisdom of its nuclear policy in spite of criticism from the House of Commons Select Committee on Energy in a report published earlier this year (see *Nature* 19 February, p.621). In a reply to that report, released last week (*Nuclear Power*, Cmnd 8317, HMSO, £2.30), the Department of Energy reiterates its belief that Britain could need 20 GW of new electricity generating capacity by the year 2000 and should go ahead with the programme, outlined in 1979, to order one new nuclear reactor a year for the next decade (15 GW in all). But it agrees with the select committee that the programme should be flexible and that each new order should be assessed on its merits.

The government is, however, sceptical of the select committee's suggestion that the £15,000 million that the programme is expected to cost might be better spent on conservation, saying that the relative benefits of investing in new plant and in conservation are too difficult to estimate. But it has accepted the recommendation that the planning margins of the Central Electricity Generating Board (CEGB), 28 per cent, should be reconsidered and has charged the Electricity Council with the task. On the contentious question of estimating future electricity demand, it says that it must rely on its own estimates and those of the Central Electricity Generating Board, the latest of which assume a growth rate in the economy of 1 per cent a year until the end of the century. Some of the estimates are due to be updated.

The select committee was particularly disturbed when it came to assess the nuclear programme on economic grounds. It shared the concern of a subsequent Monopolies and Mergers Commission report on CEGB over the way the board has estimated future capital costs from historic costs. The government agrees that this is an unreliable way of appraising future investment and, taking its cue more from the Monopolies Commission than from the select committee, has asked CEGB for improvements in its investment analysis. It has also asked for reassessments of assumptions about future coal prices and availability, power station construction times and costs and the performance of advanced gas-cooled reactors.

No change is planned, however, to the

structure of the nuclear power industry, nor does the government intend to implement the select committee's recommendations for broadening its advice on nuclear matters. But the Secretary of State is already looking at ways of attracting suitable staff to the Nuclear Installations Inspectorate.

The government is still set for a public inquiry in 1982 into the siting of Britain's first pressurized water reactor, intended to start off the new programme, at Sizewell in Suffolk. Although the precise details of the inquiry have yet to be released, Sir Frank Layfield, QC, who chaired an inquiry into local government finance in 1974, was appointed its inspector last week. The government will not, however, expect the inquiry to examine the merits of the pressurized water reactor relative to CANDU, the Canadian heavy water reactor. It rejected the select committee's recommendation for an urgent new appraisal of CANDU on the grounds that the costs of adapting it to British conditions would be too high at this stage.

Judy Redfearn

Science education funding

Pressing claims

Washington

In one of his first acts as the new president of the National Academy of Sciences, Dr Frank Press, formerly director of the Office of Science and Technology Policy under President Carter, has strongly criticized the cuts in federal support for science education that have been proposed by the Reagan Administration.

Dr Press's comments on the proposed cuts, which would virtually eliminate funding for science education from the National Science Foundation (NSF), were contained in a letter to the chairman of the House of Representatives Science and Technology Committee, Mr Don Fuqua. And the letter, together with similar complaints voiced by individuals and institutions from across the country, seems to have had an effect.

Last week, the House voted to approve a 1982 budget for NSF which included an additional \$35 million for science education in excess of the Administration's request. In the process, the Democrat-controlled body rejected by 264 votes to 152 an attempt by Republicans to remove this extra money from the NSF appropriations bill and reduce the foundation's total budget to the \$1,033 million proposed by President Reagan.

Although the House vote may prove to be critical in restoring the science education funds, there is still some way to go. Also last week the Senate appropriations committee approved a budget of \$1,044 million for NSF, which includes an additional \$20 million for science education; if accepted by the full Senate, the subsequent compromise with the

House would lie between this and \$35 million.

In his letter to Mr Fuqua, Dr Press points out that a study carried out last year for Mr Carter by NSF and the Department of Education demonstrated problems caused by the deterioration of science and engineering education that "are pervasive and complex and will require determined and concerted effort by all sectors of society for their solution". **David Dickson**

A brief respite

The *Glomar Challenger*, the work-horse of the Deep Sea Drilling Project, docked last week at Southampton in preparation for the next stage (Leg 81) of the project. The previous leg was spent in the Bay of Biscay, to determine the rifting and subsidence history of the continental margin. Leg 81 will be spent near Rockall, a few hundred miles to the west of Scotland, studying the variations of subsidence across the ocean-continent boundary.



Challenger on the crest of a wave

The Deep Sea Drilling Project is at present riding on a wave of enthusiasm for its scientific record, and has so far done well in President Reagan's financial obstacle race. *Glomar Challenger's* drilling operations are funded for at least another two years, with a 1982 budget of \$26 million. Of this, \$12 million will come from Britain, Germany, France, the Soviet Union and Japan, while the bulk of the United States contribution will be provided by the National Science Foundation.

Future plans are, however, less certain because of the debate within the foundation over the competing merits of *Glomar Explorer* — the converted ex-CIA submarine-dredger. Although vastly more expensive (its expected cost lies between \$500 million and \$1,000 million over ten years), the *Explorer* programme would be half financed by the oil industry on the strength of its considerable oil-hunting potential. The scientific issues are complex, however; and a decision from Congress on whether or not to fund initial *Explorer* development is expected in October.

Philip Campbell

EEC hormone legislation

First another report

Brussels

Progress of a sort was achieved by the European Community's agricultural ministers in Brussels last week when they discussed the problem of a total ban on hormone use in livestock breeding. They decided that Community law should ban stilbenes and thyrostatics as growth promoters in livestock breeding even though these hormones are already outlawed in all the member states.

And to serve as a figleaf for the lack of agreement on other growth promoters, it was decided to wait for a scientific report before dealing with the natural hormones oestradiol, progesterone, testosterone and the synthetic hormones trembolone and zeranol. A deadline of nine months has been set for this report but quite how it will help to solve what is essentially a political problem nobody is quite sure.

Details of how this scientific report will be compiled are also still hazy. Presumably, the same national experts who form the various scientific committees advising the European Commission on veterinary or nutritional matters will assess the present state of knowledge.

European consumer groups are planning to write to their ministers to demand that these experts are chosen for their impartiality. In fact, anyone nominated by a national government will have already contributed towards that government's position on hormones. The scientific report may therefore end up reflecting only the existing divisions of opinion.

The United Kingdom and Ireland are the strongest opponents of a blanket ban, but producers in other member states are lobbying effectively in the other direction. France, Denmark and Italy want legislation to be as severe as possible — and it was the ruling of an Italian court to ban veal sales which started the whole affair. The new French agriculture minister, Edith Cresson, stressed at the Council that France would not be bound by the conclusions of the scientific report. Belgium and Germany in theory already ban the use of all hormones.

Consumer groups are also incensed by the ministers' reluctance to thrash out the ways and means to enforce a ban of any kind. Natural hormones are notoriously difficult to detect and the various member states do not all have the same reputation for the effective policing of detectable substances.

What has been achieved is the right for any member country to prevent the entry from another EEC country of meat which contains hormones such as stilboestrol, diethylstilboestrol, dienoestrol and hexoestrol. And those member states with legislation which bans other hormones can carry on as before.

In nine months' time, the European

Community will still have to decide between a policy which favours the consumer approach of banning unless proved safe and one favouring the industry's attitude of carrying on until a substance is known to be dangerous to human health — which is the choice of the United States.

Jasper Becker

Third World development

Proper priorities

Washington

The US Agency for International Development (AID) is continuing to pick up the pieces following the failure of the Carter Administration to convince Congress of the need for a new Institute for Scientific and Technological Cooperation to act as a focus for research on the needs of Third World countries.

Earlier this month the agency announced that as part of a general reorganization, a new Bureau of Science and Technology was to be established. Furthermore the title of "assistant administrator for development support" will be changed to "senior assistant administrator for science and technology", a position that will have responsibility for the bureau.

One of the aims of this reorganization, according to agency officials, is to strengthen the research activities of AID which have in the past been of lower priority than more conventional technical aid programmes. This was one of the reasons quoted for establishing the Institute for Scientific and Technological Cooperation, a body which would remain under the same umbrella as AID, but operate independently with its own board of advisers.

The parallel between the goals of the ill-fated institute and the newly proposed bureau is strengthened by the fact that Mr Reagan's appointee to head the bureau, Dr Nyle Brady, had previously been successfully wooed by the Carter Administration as its proposed director for the institute. Dr Brady will have responsibility for the administration of a wide range of centrally-funded research and development programmes, some of which were to have been shifted from AID to the new institute.

The Carter Administration had originally proposed significant increases in research support for these areas which would have virtually doubled the AID budget for research and development to \$229.8 million in the fiscal year 1982. The Reagan Administration has reduced this dramatically to a request for \$134.5 million, with large cuts, for example, in support for agricultural research on food production in Africa.

Even with these cuts, however, AID officials point out that the Administration's request for science and technology programmes for 1982 is 40 per cent higher than the amount spent on

research by AID in 1980. A particularly significant increase is in energy research, scheduled to grow from \$3.4 million in 1980 to \$16.3 million in 1982.

The impact the increase in funds and the internal reorganization will have remains controversial. Those in Congress who argued most strongly for the institute are somewhat sceptical.

This scepticism is reflected in a report from the House Foreign Affairs Committee on the AID budget request, which has included as a separate item a proposed \$10 million (reduced from President Carter's requested \$18 million) for the support of scientific and technological cooperation with developing countries. The committee has reinserted this money into the broader category of "selected development activities".

In fact, much of this money will be spent in a way comparable to what the institute would have done, but through a different agency, the National Academy of Sciences. AID has agreed to provide \$35 million of its science and technology funds over the next five years to a new grants programme established under the academy's Board on Science and Technology for International Development.

The grants, which will be \$5 million for 1981 and increase to \$10 million by the middle of the decade, will be spent largely on supporting research activities in the developing countries on projects selected by an advisory board. "One of our principal goals is to help developing countries build up their research capabilities, precisely the request that they were making at UNCSTD (the United National Conference on Science and Technology for Development) in Vienna in 1979", says the board's staff director, Dr Victor Rabinowitch.

When the grants programme is added to its more traditional activities, such as arranging seminars on Third World problems, the board will begin to resemble a scaled-down version of the Institute for Scientific and Technological Cooperation.

Furthermore its relative autonomy from AID goes some way to meeting the objections of those who argued that, as put forward by the Carter Administration, the inclusion of the institute under the umbrella of development aid bodies would have stifled its flexibility. **David Dickson**

Polish economic reform

Self-governance wins

Warsaw

The latest Polish plans for economic reform, discussed in considerable detail during the recent extraordinary Party congress, include a major shake-up for Polish science and technology. The basis of the proposed new economic model is a switch from rigid central planning of industrial production to a degree of "self-governance" (*samorządność*) — a word introduced into the Polish political lexicon

with the rise of the independent trade union movement last August.

The new system, which is to be introduced in four phases over the next 18 months, will allow for more public and expert discussion in advance of overall production plans and delegate to factories initiative for deciding how they fit in with the overall targets.

As in Hungary, the authorities will steer production by a system of financial incentives, but factory managements will have considerable responsibility for decision-making on modernization and technological innovation. One of the major charges laid against the previous Gierek government in the preamble to the new plan is that autocratic planning blocked technical innovation. At the same time, many of Mr Gierek's prestigious schemes are being dropped. The "Program-Wisla", for example, which was totally to remodel Poland's major waterway by the end of the century, has been cut back to its environmental and agricultural components. The schemes for hydroelectric dams and the straightening of rivers have been dropped *sine die*.

Other harm done by Gierek's policies will not be so easily remedied. Discrimination against the private farmer in the pricing and supply system ran agriculture down to a level where even the new rationing system cannot ensure supplies of basic home-produced foods. The abolition of the small pharmaceutical cooperative factories in the 1970s has meant that Poland has had to shop abroad for a wide variety of drugs whose production runs were too small for the state pharmaceutical enterprise Polfa — and when hard currency ran out, so did the supply of even the most essential drugs. The international relief operation to ensure vital supplies now being mounted jointly by Solidarity and the Polish Ministry of Health is a strange contrast with the anxiety in the mid-1960s in the United Kingdom that Polish drugs might undercut the domestic products.

Although the new plan pays lip-service to the strengthening of scientific and technical ties with other (especially Comecon) countries, it is also clear that for the present Poland will have to rely mainly on home-based technology.

A reorganization of the responsible ministries is also on the cards. One is for a State Committee for Technical Progress responsible to the Prime Minister. The new body would be unique in the Comecon bloc in dealing only with technology, and may point to an ultimate break-up of the Ministry of Science, Higher Education and Technology. Straws in the wind include the appointment last month of a mining specialist, Jerzy Nawrocki, as Minister of Science, Higher Education and Technology, and the plans discussed at the congress for a complete review of the country's industrial research institutes and the closure of those not cost effective.

Whatever the final form of the reorgani-

zation, one change is certain — a revision of the present practice of financing science by "problems" ostensibly related to economic needs and organized into a cumbersome hierarchy of priorities. Among likely innovations are direct research contracts between industry and the universities or academy institutes and the allocation of funds for basic research directly to the institutions concerned.

For the moment, however, virtually all such funding will come in non-convertible zloty. Hard currency for the purchase of equipment, reagents or journals from the West is simply not to be had; according to one leading physicist, only two subscriptions to *Nature* are now authorized for the whole of Warsaw. **Vera Rich**

British Technology Group

Mixed marriage

Last week, Britain's National Enterprise Board and the National Research Development Corporation finally merged. The British Technology Group, as the two are now called, has been in the offing since Sir Frederick Wood was appointed chairman of both the board and the corporation at the beginning of the year. His plan for a merger has now been accepted by the government, although legislation to give the group formal status will not be possible until 1983.

The plan is that the group will offer more streamlined services to innovators seeking backing for their ventures than either of the parent organizations. The merger will also eliminate competition between the board and corporation for the same projects.

The function of the former corporation to support research and development into potentially marketable innovations will be largely taken over by the group's technology transfer division. And the board's experience in raising finance to invest in industry will be incorporated into the new investment and operations divisions, the latter being responsible for returning investments to the private sector.

Both parent organizations have been widely criticized in the past. The board, which was originally set up in 1965 by the then Labour government with the potent title of Industrial Reconstruction Corporation, has since been substantially curtailed. And the corporation has been criticized mainly by academics for not being adventurous enough in supporting their innovations and persuading industry to take them up. The British Technology Group is intended primarily to marry the corporation's access to innovators with the board's financial expertise. There is to be a new liaison team whose prime responsibility will be to scour the universities more aggressively than in the past.

The group is also to have a new corporate development unit to develop strategies for investment. Areas of priority are advanced manufacturing and biotechnology. A

study team on robotics has already presented a strategy paper to the group which is expected to result in investments within the next few months. The group also has plans for investing in biotechnology in areas not covered by Celltech, the company set up by the board last year.

As yet the group does not know precisely how much public money it will have to spend. The precise structure of the new organization should become clearer in September.

Judy Redfearn

Tidal energy

The big dam

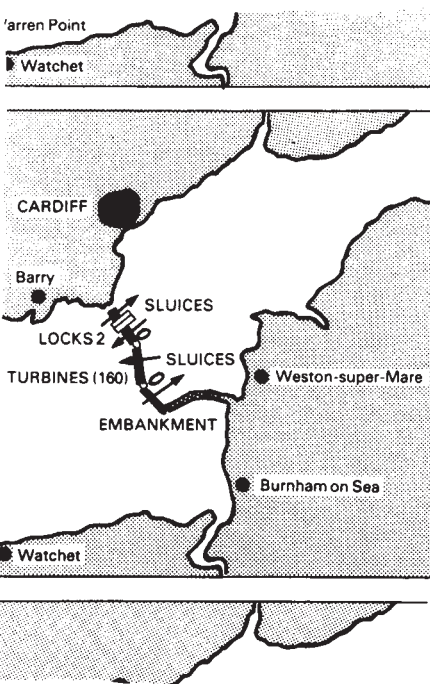
The economic feasibility of a tidal barrage for generating electricity across the Severn estuary in south-west England will depend on how many nuclear power stations are attached to the national grid. In most circumstances, however, a barrage would be a good investment, according to a report by the Severn Barrage Committee published last week. The committee, set up in 1978 under Sir Hermann Bondi, then chief scientist at the Department of Energy, and whose deliberations have cost £2.3 million, has also concluded that a barrage is technically feasible but has called for detailed studies of its potential environmental, industrial and social impact.

The outlook for a barrage across the Severn estuary, which has one of the greatest tidal ranges in the world, has improved considerably in the past two years, according to the committee. Approximately 40 per cent more energy should be obtainable from an inner barrage, the scheme favoured by the committee, than was originally thought. And electricity generation should be possible nine rather than 16 years after the start of construction.

The committee estimates that the most cost-effective scheme would cost about £5,600 million and would generate 13 TWh a year from an installed capacity of 7,200 MW. That scheme is for an inner barrage from Brean Down to Lavernock Point consisting of large prefabricated concrete units housing turbines and sluices. A more ambitious barrage built further out to sea, costing £8,900 million and generating 20 TWh from 12,000 MW capacity, would be less economic and involve greater risk during construction. The inner barrage scheme, on which the committee has based most of its calculations, has the added advantage that a second barrage could be added at a later stage. The two together would provide as much energy as the outer barrage at a similar overall cost.

Because of the variability in tidal energy, a barrage would reduce the need for other generating capacity only by about 1,000 GW, according to the committee, most savings being in fossil fuel. And as nuclear power is also capital-intensive and low on fuel costs, the attractiveness of a barrage will to a large extent depend on future plans

for nuclear power. Under the Central Electricity Generating Board scenario, which assumes a fairly rapid growth in nuclear power stations, the barrage is marginally unattractive, says the committee. But it could provide an attractive alternative if nuclear plans are modified only slightly.



The committee seems convinced that a Severn barrage should at least be seriously considered as a possible alternative to investment in nuclear power. Before a decision is finally taken, however, it recommends further studies of the effects on water levels, wildlife and local amenities and the capacity of British industry to cope with such a project. If the government decides to take investigations further, the next step should be an acceptability and preliminary design study costing £20 million and the building of prototype concrete units at a cost of £25 million.

Judy Redfearn

US nuclear weapons

Lasers purify

Washington

A Washington environmentalist group is trying to prevent the Department of Energy from carrying out an expanded programme of research into laser isotope techniques for separating plutonium for US nuclear weapons from spent nuclear fuel.

A classified research programme into laser isotope separations, using copper vapour lasers for uranium enrichment, has been carried out at a relatively low level of funding at the Department of Energy's Lawrence Livermore National Laboratory and the Los Alamos National Laboratory since the mid-1970s — in both 1980 and 1981 total funding for the programme was \$5 million.

In testimony before a congressional committee in March which has just been declassified, the Department of Energy requested an increase in funds to \$155 million over the next three years to support this research and its application to plutonium enrichment. This money would include the construction of a pilot plant, and the department is also discussing plans to construct a \$200 million multipurpose plutonium isotope production plant.

In a letter sent last week to the Secretary of Energy, Mr James Edwards, a staff scientist and an attorney with the Natural Resources Defense Council say that the plutonium laser isotope separation programme "carries grave nuclear weapons proliferation risks" and is "totally inconsistent with the fundamental objective of President Reagan's nuclear non-proliferation policy".

They point out that although the department has identified several other applications for the process — such as the enrichment of fuel-grade plutonium as fuel for the breeder reactor research programme — "it appears that these applications would not serve as adequate justification for the program, particularly in light of the severe proliferation dangers".

In their presentation to the congressional committee, Department of Energy officials said that the purpose of the expanded programme was to make available for weapons use approximately 70 tonnes of plutonium by reprocessing commercial spent fuel and then enriching the plutonium to weapons-grade by laser isotope separation. It would be possible to convert fuel-grade plutonium produced by the department's nuclear reactor in Hanford, Washington, to weapons-grade.

One of the purposes of the laser isotope separation process is to enrich plutonium to weapons grade by reducing the concentration of the major contaminant isotope Pu-240, which is not spontaneously fissile. According to testimony given to Congress by F. Charles Gilbert, assistant secretary for nuclear materials at the Department of Energy, plutonium is placed in one of four categories depending on the concentration of the Pu-240 contaminant: reactor-grade for 19 per cent or greater, fuel-grade for between 7 and 19 per cent, weapons-grade for less than 7 per cent, and supergrade for between 2 and 3 per cent.

The laser enrichment process has two non-military uses. The first is the separation of the Pu-238 isotope, which is used in space satellites and heart pacemakers, as well as for a heat source in thermoelectric generators. The second is the separation of the Pu-241 isotope which decays to Am-241 which is used in smoke detectors and for logging wells. According to the Department of Energy, the non-weapons applications at present require less than two per cent of the output of the Savannah River plant, which is the current source of supergrade plutonium.

David Dickson

CORRESPONDENCE

Man versus tsetse

SIR — I must contend with your leading article "Fighting tsetse flies and people" (*Nature* 2 July, p.1).

In it you state that the eradication of the tsetse fly in certain parts of Africa (by such organizations as ILRAD) would be a revolutionary advancement for its people. Cattle would then be able to graze over a much larger area of land, so more food would result.

This, in itself, is quite true — for man. But the tsetse fly is one of the last remaining barriers against mankind's inevitable destruction of the natural habitats for literally millions of animals of the continent. Would not this "advancement" seal the fate of these millions?

Animals in the wild cannot compete with organized grazing. It's like asking one man to stop a moving steamroller by standing in front of it — he would be crushed. The question is: do we have the right to do this, to crush other forms of life?

As the most powerful animal on this planet, man is entrusted with a duty to preserve that which cannot defend itself. Although the African lion is strong, it can't stop a steamroller.

Jamesville,
New York, USA

ADAM J. BECKER

Author — Anon

SIR — Drs Gillman and Lichtigfeld (*Nature* 25 June, p.608) suggest that the present system of assessment of scientific papers is endangered by a cabal of corrupt referees, and propose a completely open system. Referees may be prejudiced and incompetent but they are not a secret society. Vigilant editorial control rather than public exposure seems the most appropriate way to eliminate bad referees.

Since referees may be influenced by the known standing of the author or the author's institute, I propose that the anonymity afforded to referees should be extended to authors. If papers were submitted with the author's name, address and acknowledgements on separate sheets of paper and these were retained by the editor when the manuscript was forwarded to the referees, conscious or unconscious bias would be avoided.

P.M. GAYLARDE

Royal Free Hospital,
London NW3, UK

From the museum

SIR — We should like to explain a matter that puzzled you in your editorial of 2 July (p.1): the place of cladistics in the exhibition policy of the British Museum (Natural History).

Exhibitions have a general tendency to conceal background controversy¹. This should not be seen as a conspiracy on the part of designers but is, in fact, a phenomenon that is apparent in the majority of educational media. Every lecturer or textbook writer has to decide the extent to which he will emphasize a straightforward story at the expense of alternative views. The emphasis chosen depends on the circumstances; the less

knowledgeable the audience, the less appropriate it is to elaborate the story. Knowing how to make this choice is particularly difficult in exhibitions since the limits of educational attainment are even less well understood than in other more generally practised media. In public museums there is an overriding need "to deliver an interesting and stimulating message to laymen"².

One of the essentials for good communication in a museum is to set objects within some kind of context so that specimens take their place in a coherent line of thought and do not seem to be a series of unrelated items. Without such a context, the specimens have little meaning and odd facts that are gleaned tend to be quickly forgotten; in educational terms, the exhibit would be a failure.

In two of our recent exhibitions, Man's place in evolution and Dinosaurs and their living relatives, the context for the specimens has been that of phylogeny and the fact that we chose to present phylogeny in cladistic terms has been seen by some critics as an attempt to promote cladistics at the expense of more traditional views. Recent correspondence in *Nature* and elsewhere emphasizes the fact that a heated debate is in progress about the relative merits of cladistic and traditional approaches to phylogeny and systematics. Exhibitions have no part in this debate between scientists, but it is clear that whichever system is used in an exhibition, one faction or another will be provoked into stern response. Accepting that one of the competing systems had to be chosen, we judged on educational grounds that cladistics was the most appropriate because of its clear and overt reasoning. Once a layman has been introduced to the system he is in a better position to understand why and how the scientists reach their conclusions. No one would pretend that the use of a cladistic system makes the study of phylogeny easy, but it can make the steps in the argument more apparent.

Some critics feel that the exhibitions should have said more about the existence of alternative systems and we are making some changes to meet that criticism, but the danger is that the exhibits in question may become displays about phylogenetic systems rather than about fossil man, or dinosaurs. A separate exhibition about systems of classification which is scheduled for future display may help to give a fuller presentation of current views.

The recent controversy has highlighted how difficult it is to maintain a balance in exhibitions between directness and fairness to every point of view — exhaustive can be exhausting to the visitor. The application of educational principles to museum exhibitions is a subject that has been sadly neglected in the past and the credibility of museums as a means of informal education has suffered in consequence. We should like to encourage serious discussion and research in this field.

R.S. MILES
G.C.S. CLARKE

Department of Public Service,
British Museum (Natural History),
London, SW7, UK

1. Lewis, B.N. *Mus. J.* 80, 151-155 (1980).
2. Wittlin, A.S. *Curator* 14, 138-150 (1971).

Suiting everyone

SIR — You object most strongly (*Nature* 12 March, p.75) to the phrase "If the theory of evolution is true . . .". Having myself tried more than once to explain the work of a scientific institute in everyday language, I have some sympathy with the author of the museum guide. When all is said and done, scientific knowledge is provisional and fallible, as, implicitly, you agree. Your objection seems to be that this particular form of words gives a handle to those opponents of evolutionism whose views you regard as not "respectable", specifically to creationists. I am not a creationist and accept the evidence that evolution by natural selection has taken place, ultimately producing Man. Nevertheless I sometimes find biologists, when writing for a popular readership, to be overly dogmatic on this point. The creationist movement on this continent (which is not simply a revival of fundamentalism and obscurantism) is partly inspired, I believe, by a healthy reaction to this dogmatism.

To represent our scientific theories as final when they are not is in the long run to do our own cause a disservice, and is, in any case, dishonest. Can any evolutionist provide a clinching argument that chance mutations acted on by natural selection *alone* are the mechanism of evolution? I think not, and, indeed, I myself find it at least as easy to believe that there has been a purposive element in evolution. Did you really mean to imply that it is not "respectable" to believe "that the course of events may be determined by literally supernatural influences"?

In your last paragraph, you almost seem to be proposing a *noble myth*, the purpose of which will be to keep all us scientists working happily at our research. If we recognize too clearly the logical status of our theories, you argue, the scientific enterprise will be undermined. But Popper's ideas, as you admit, have been around since 1934, and many branches of science have been uncommonly vigorous since then. Your argument has something in common with the old discredited argument that Christianity is socially beneficial and should be taught to the young whether it is true or false. It is this kind of argument that plays into the hands of creationists, by making more plausible their claim that evolutionism is not a scientific theory, but a kind of religion.

The most helpful ideas in this situation, I find, are not Popper's, nor dogmatism, but Kuhn's. Almost all of modern biology is "normal science" within the paradigm of Darwinism-Mendelism. Always there are some facts that do not fit a paradigm, and one day they may become so bothersome we will have to change it. In the meantime, it provides a frame of reference in which we work and by which we make sense of a wide range of data and experience. Darwinism-Mendelism is the best paradigm we have (in its field) and, as Darwin said, gives grandeur to our view of life, but it should not fight its critics by claiming, like them, to be based on divine revelation.

A.H. BATTEN

National Research Council of Canada,
Victoria, British Columbia, Canada

Too tolerant

SIR — I agree with you (*Nature* 28 May, p.271) that tolerance of religious fundamentalism in the Natural History Museum has gone too far. But you underestimate the menace to science, and to the propagation of rational thinking by education, presented by this tolerance. I see two problems — first the conflict between science, in the broadest sense of the word, and mythology. The second is the egalitarian and anti-elitist attitude prevailing among many educated people which has brought about the woolly-minded tolerance you describe.

Homo sapiens indulges in two main kinds of thinking about his world which must go back to the very emergence of the first intelligent, tool-making primates. The appearance of the first stone tools, adapted for specific uses, implies the existence of empirical thinking involving designing an implement for future use, testing it practically and modifying it in the light of experience. From this technological thinking, based on trial and error, must ultimately have developed scientific thinking, based on observation, hypothesis, prediction, experimental test and modification.

On the other hand myth-making must have emerged at an equally early time to allow primitive men and women to explain the vast range of often frightening phenomena outside their understanding.

We might also describe this duality of thought processes in terms of deductive and inductive thinking. In the former the grand, explanatory ideas come first (usually being taught as dogma) and are then used to interpret everyday natural and social phenomena. In the latter the explanatory general hypotheses are developed specifically to explain a body of facts or phenomena. I am now using the term "myth" to describe all preconceived ideas or beliefs reached deductively (before examining the evidence properly) and "science" to describe the opposing, inductive method of explaining things (which also involves testing the hypotheses devised).

The conflict between these two modes of thought seems to be as basic now as it was in Darwin's time and to have extended into many more fields. Myths are believed in for a variety of reasons, both reputable and otherwise, but it seems reasonable to suppose that they reflect psychological or emotional needs. Their great variation suggests this, the zeal with which they are believed being the only common factor. The dedicated communist, Christian, Buddhist, believer in flying saucers or in a tribal god each believes his creed deductively (usually having been taught it *a priori*) and each is equally convinced of its truth, and of the folly or evil of those who disagree. The only common factor seems to be a strong, often overwhelming, perhaps unconscious desire to believe in some set of ideas which explain phenomena one is interested in or worried about. This is the exact opposite of the scientific method.

It is essential to realise that what is being described here is a mental attitude and not categories of rational and irrational subjects to think about. Of course it is possible to be a communist or a Christian on inductive, rational grounds if one has studied as much of the evidence as possible, considered alternative explanations for it and found them less convincing. Likewise it is unfortunately true

that a scientist can elevate the status of his favourite theory in his own mind to that of a myth — and believe it implicitly, emotionally and against all opposition. It is the mental attitude, not the subject matter, which produces both myth and science.

It follows that there can be no place for myths either in science or in any proper education programme that tries to avoid indoctrination. "Creation science" is a myth which depends on a dogmatic and deductive faith in the existence of a God who has intervened frequently on Earth to make new living things out of nothing and, this being so, the Natural History Museum may be said to have betrayed both its scientific status and its educational role by implying that this myth is somehow on the same intellectual plane as the theory of evolution.

Where will this process end? Another interesting if minor example of a new myth about human origins appeared in *The Sunday Standard* of 31 May. A female body builder, when explaining why she pursued this hobby, is reported as saying "We were never meant to be the weaker sex, and being seen as such is a fairly recent evolutionary development. Originally it was women who went out hunting to find food; they were the strength of the family." This relatively harmless myth, though it bears no relation to the facts of archaeology and anthropology, may be assumed to serve a useful purpose to the lady concerned by giving her the comfort of an ideological justification for what most people probably regard as a bizarre hobby. But to the Natural History Museum it must surely now be a valid alternative hypothesis, or even a theory, about the early development of man and society. Perhaps the Public Services Department will consider adding a new panel on "Ms Cheshire's hypothesis" to its display

on human evolution? Or is it the popularity of a myth which qualifies it for serious attention in Cromwell Road now? If so then a panel describing Von Daniken's ideas about the bringing of civilization to Earth by spacemen in prehistoric times is surely required.

My explanation for this excessive tolerance, which is an irresponsible failure to distinguish between scientific thinking and irrational myth-making, is that it is an effect of modern egalitarian thinking. This is quite widespread among some university staff, extending to advocating equal tolerance for the "views" of all people, no matter how well or ill-qualified in the subject concerned. I have actually heard it argued in faculty meetings in this university that it is no longer necessary to have professors as more than administrative departmental heads, to have them leading and guiding their staff in research and academic standards by their example; everyone, it seems, should now be "capable of standing on his own feet". Is this phenomenon sometimes due to a real fear of being thought elitist, "better" than one's fellows, and to a failure to realise that better trained or informed is not the same thing as absolutely better?

It cannot be a coincidence that no equivalents of the "creation science" myth appear in disciplines like engineering or navigation where equal tolerance for them would produce catastrophe. The tolerance of myths about human origins by the staff of the Natural History Museum will not cause their building to fall down about their ears but it has certainly undermined the institution's reputation as an authoritative laymen's guide to the world of science.

E.W. MACKIE

Hunterian Museum,
The University of Glasgow, UK

Halstead's defence against irrelevancy

SIR — In his Presidential Address to the British Association in 1977, Sir Andrew Huxley "wondered what topic could, in the second half of the twentieth century, generate emotions as strong as those which arose over evolution in the early 1860s". If the recent correspondence in the columns of *Nature* is anything to go by then it undoubtedly remains this selfsame topic.

Previously I drew attention to what was taking place at the British Museum (Natural History) which occasioned serious disquiet in the scientific community. I suggested a possible explanation, but whether this is rejected or accepted does not affect the substantive issues I raised. So far, no serious attempt has been made by anyone in authority to respond to the accusations that have been levelled.

The present debate began in 1978, when the trend of the new policies of the museum's Public Services Department and the proposed presentation of dinosaurs merely to illustrate cladistics were discussed at a symposium at Reading (see refs 1,2). In 1979 the exhibit "Dinosaurs and their living relatives" opened and the accompanying booklet spelt out clearly exactly what cladistics involved. The crude didactic method of its presentation contrasted dramatically with the popular but nonetheless scholarly approach of the museum's book *"A New Look at Dinosaurs"* by Alan Charig, published at the same time³.

Then in 1980 came the exhibit "Man's place

in evolution", again with its accompanying booklet in which it was again emphasized that we can never know how species arose in the past and that we assume that new species arise only by splitting from previous species. But the booklet includes the most amazing assertion of all, that no fossil species can be considered the direct ancestor of any other. I referred⁴ to the place of *Homo erectus* and the Petrolona skull which sits on the morphological boundary between *Homo erectus* and archaic *Homo sapiens*. The review article on fossil man by Cronin *et al.*⁵ and also recent work on planktonic organisms by Prothero and Lazarus⁶ have dealt with this matter effectively. The museum has now, in 1981, removed the labels which previously stated categorically that the *Homo erectus* people were not our direct ancestors.

It is a fact that the dinosaurs are used to illustrate the principles of cladistics and not their natural history. Similarly, the evolution of man is being used as a vehicle for cladistics, even though this has involved crudely distorting the evidence. It is worth recalling the words that Andrew Huxley used in his Presidential Address to the British Association in 1977:

"Scientists whose findings contradict fashionable social theories are accused of distorting their results through political prejudice and it is suggested that on sensitive topics we should base our beliefs not on what is actually found to be the case, but on the supposed consequences of holding particular beliefs — in effect

that we ought to replace science by wishful thinking. I regard any such attempt to deflect scientific conclusions for political or social motives, however well-meaning, as a betrayal of science."

The situation at the British Museum (Natural History) is suspiciously comparable to that against which Huxley's criticisms were directed.

It was, however, something even more astonishing than all this that stung me, a palaeontologist and evolutionist, into activity⁷. This was the assurance given in writing in 1978 by the Head of the Public Services Department to the creationists⁸ that the museum in its new exhibition on evolution (to be opened in 1981) would make it quite clear that the theory of evolution was not a scientific theory. It transpired that this attitude stemmed from the writings of Sir Karl Popper. The museum was now concerned with "concepts of science" as adumbrated by Popper, and the articles of Platnick and Gaffney⁹ and even more recently Tassy¹⁰ emphasized the scientific (in the sense of Popper) nature of cladistics. In marked contrast, palaeontology, the historical side of evolution, was dismissed because it comprised unique events which were by definition unrepeatable and so not subject to test and hence not part of science. This is the view of Patterson¹¹ which he claimed comes from Popper.

The notion that the process of evolution by natural selection is non-science was based on the idea that "survival of the fittest" is a tautology and hence unfalsifiable. On this point Popper¹² wrote:

"The fact that the theory of natural selection is difficult to test has led some people, anti-Darwinists, to claim that it is a tautology . . . Since the explanatory power of a tautology is obviously zero, something must be wrong here . . .

"I mention this problem because I too belong among the culprits . . . I have changed my mind about the testability and the logical status of the theory of natural selection; and I am glad to have an opportunity to make a recantation."

Ruse¹³ and Flew¹⁴ have also dealt with this. (The museum, although initially keeping faith with the creationists, has since withdrawn the film loop concerned, following the adverse review of the exhibition by Barry Cox¹⁵ (see also Miles¹⁶).

Popper¹⁷, following my attack⁷ on the ideas being peddled in his name, has emphatically dissociated himself from his self-proclaimed disciples in the following terms:

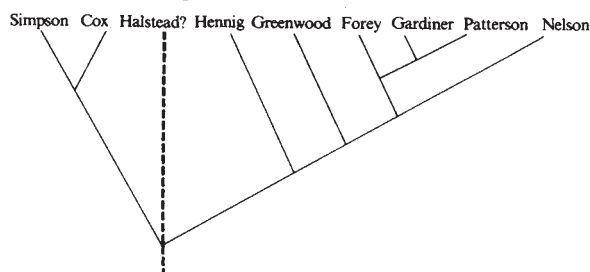
"It does appear that some people think that I denied scientific character to the historical sciences, such as palaeontology, or the history of the evolution of life on Earth. This is a mistake, and I here wish to affirm that these and other historical sciences have in my opinion scientific character; their hypotheses can in many cases be tested. It appears as if some people would think that the historical sciences are untestable because they describe unique events. However, the description of unique events can very often be tested by deriving from them testable predictions or retrodictions."

With the philosophical aspect in better perspective, it became possible to consider what might lie behind the attitudes to evolution and palaeontology being unveiled by the museum's Public Services Department. My suggestions on this stimulated a certain degree of controversy.

Sir Andrew Huxley in his opening of the museum's new exhibit "Origin of Species" stated that he had "been quite unable to comprehend the suggestion that cladism is

How to classify the cladists

Sir — May I propose a cladogram of cladists?



The cladogram is, of course, based only on derived characters appearing late in ontogeny; it makes no claims as to ancestry. (Apologies to Halstead, but he does display a confusing array of derived traits.)

London, UK

BRIAN LEITH

somehow antagonistic to evolution, or that cladism is linked to the theory that evolution progresses by fits and starts, or that cladism is more Marxist than other styles of classification".

That Hennigian or classical cladistics is antagonistic to the gradualist type of evolution is surely evidenced by Rosen, Nelson and Patterson¹⁸:

"Hennig established a criterion of demarcation between sciences and metaphysics at a time when neo-Darwinism had attained a sort of metaphysical pinnacle by imposing a burden of subjectivity and tautology on nature's observable hierarchy. Encumbered with vague and slippery ideas about adaptation, fitness, biological species and natural selection, neo-Darwinism (summed up in the "evolutionary" systematics of Mayr and Simpson) not only lacked a definable investigatory method, but came to depend, both for evolutionary interpretation and classification, on consensus or authority."

This sounds antagonistic to me.

The latest brand of cladism, the transformed variety, has undoubtedly slid into what seems to me to be an overtly anti-evolutionist stand. Patterson¹⁹ writes:

"Hennig's 1966 book, as the title *Phylogenetic Systematics* suggests, was based in evolutionary theory. But as the theory of cladistics has developed, it has been realised that more and more of the evolutionary framework is inessential and may be dropped. Platnick (1980) refers to the new theory as 'transformed cladistics' and the transformation is away from dependence on evolutionary theory. In my view, the most important outcome of cladistics is that a simple, even naive method of discovering the group of systematics — what used to be called, the natural system — has led some of us to realise that much of today's explanation of nature, in terms of neo-Darwinism, or the synthetic theory, may be empty rhetoric. Eldredge and Cracraft (1980) provide a detailed criticism of the transformational approach in biology, and the ensuing interest in storytelling about adaptive change."

The link between cladistics and the theory that evolution progresses by fits and starts

(punctuated equilibria) is emphasized in particular by Cracraft²⁰ and is certainly the type of association that a normal reader would obtain from reading Eldredge and Cracraft's *Phylogenetic Patterns and the Evolutionary Process*²¹.

The concept of punctuated equilibria is undoubtedly linked to Marxism, according to Gould²², who claims to be a Marxist. Gould²² sees "notions of gradualism arising largely out of a pervasive political bias [and] . . . the replacing of gradualism with the flip-like style of change which has been appreciated within Marxist philosophy for a long time."

The current drive against the concept of gradualism is motivated primarily by Marxists — this does not mean that there are not plenty of non-Marxists who agree nor does it mean that the views may not turn out to be true. What is objectionable is the distortion of scientific data for ideological purposes.

Finally, I am not opposed to either a cladistic or a Marxist approach being presented. My objection is to the attempt to impose highly controversial concepts on to the general public not by argument or discussion but simply by unsubstantiated assertion. Whalley and Gibbon²³ have pointed out that the current policies being pursued by the museum are diametrically opposed to the 1972 paper to the trustees, on which they are supposedly based, where it was emphasized that the new exhibition scheme "should point out areas of doubt and speculation".

Whatever view is taken of my linking hypothesis, this does not in any way affect the factual account of what has been happening in the museum. It is to be hoped that the minor changes made in the past few weeks herald a new awareness by the trustees of their responsibility towards maintaining integrity in the natural sciences.

L. BEVERLY HALSTEAD

University of Reading, UK

1. Halstead, L.B. *Nature* **275**, 683 (1978).
2. Halstead, L.B. *Nature* **276**, 759–760 (1978).
3. Chaz, A. *A New Look at Dinosaurs* (British Museum (Natural History), London, 1979).
4. Halstead, L.B. *Nature* **288**, 208 (1980).
5. Cronin, J.E., Boaz, N.T., Stringer, C.B. & Rak, Y. *Nature* **292**, 113–122 (1981).
6. Prothero, D.R. & Lazarus, D.B. *Syst. Zool.* **28** (1978).
7. Halstead, L.B. *New Scientist* **87**, 215–217 (1980).
8. Miles, R.S. *Creation News Sheet* **28** (1978).
9. Platnick, N.I. & Gaffney, E.S. *Syst. Zool.* **26**, 360–365 (1977); **27**, 137–141, 381–388 (1978).
10. Tassy, P. *Int. Symp. Conceptual Methods in Palaeontology*, Barcelona, 65–73 (1981).
11. Patterson, C. *Evolution* (British Museum (Natural History), London, 1978).
12. Popper, K. *Dialectica* **32**, 344 (1978).
13. Ruse, M. *Phil. Sci.* **44**, 638–661 (1977); *New Scientist* **90**, 828–830 (1981).
14. Flew, A. *A Rational Animal* (Clarendon, Oxford, 1978).
15. Cox, B. *Nature* **291**, 373 (1981).
16. Miles, R.S. *Nature* **291**, 530 (1981).
17. Popper, K. *New Scientist* **87**, 611 (1980).
18. Rosen, D.E., Nelson, G. & Patterson, C. in *Phylogenetic Systematics* (University of Illinois Press, 1979).
19. Patterson, C. *Biologist* **27**, 234–240 (1980).
20. Cracraft, J. in *Phylogenetic Analysis and Palaeontology* (Columbia University Press, 1979).
21. Eldredge, N. & Cracraft, J. *Phylogenetic Patterns and the Process of Evolution* (Columbia University Press, 1980).
22. Gould, S.J. *Sci. Nature* **3**, 5–12 (1979).
23. Whalley, R. & Gibbon, K. *Palaeont. Ass. Circ.* **105**, 4 (1981).

NEWS AND VIEWS

Thalassaemia: from theory to practice

from Bob Williamson

THE international symposium on recent advances in thalassaemia, held in Sardinia from June 7th to 11th, was divided quite sharply into two parts. Several molecular biologists presented the latest data on the gene defects causing the various types of thalassaemia found throughout the world, following which doctors from the countries where the disease is common exchanged experiences on the use of this knowledge in programmes for population screening and antenatal diagnosis. The meeting was the first in three years to involve most of the molecular biologists and all of the leading figures involved in control of haemoglobinopathies, and allowed some interesting conclusions to be drawn.

Since the first demonstration that most cases of α -thalassaemia (with depressed α -globin synthesis) are due to gene deletion, the exact gene defects in each of the haemoglobinopathies have been analysed one after another. As the thalassaemias, and in particular β -thalassaemia (in which the synthesis of the β -globin chain synthesis is impaired), protect carriers against malaria, it is not surprising that many independent mutations have spread in populations in malarial areas. One interesting possibility put forward by Herman Lehmann (University of Cambridge) is that α -thalassaemia might not really protect against malaria at all, but rather mitigates the effects of β -thalassaemia (or sickle cell disease) and so is often found in association with it.

Genetic engineering techniques have now allowed a complete understanding of the causes of the major forms of β -thalassaemia in Italy, Greece and Cyprus. In β^0 -thalassaemia, A. Bank (Columbia University) showed that a point mutation has occurred in many patients at the 5' intron-exon junction of the large intron so that splicing of the nuclear RNA transcript cannot take place at all. Patients homozygous for this point mutation cannot make messenger RNA, and therefore make no β -globin. Fortunately, the splice junction is also a restriction site for the enzyme *HphI*, and therefore a direct gene analysis (of the patients, or antenatally) should be relatively simple.

In the case of β^+ -thalassaemia, the form most common in Cyprus (and therefore in London), B. Forget (Yale University) reported results from his laboratory and from that of R. Williamson (St Mary's, London) which show that the mutation is within the small intron and creates a new splice site which can act as a preferred alternative to the correct one. This explains why patients are still able to make a small amount of β -globin. Yet another type of mutation, causing premature termination of the β -globin protein chain, has been identified by Y.W. Kan (University of California, San Francisco). It is now clear that the majority of β -thalassaemia cases, as well as those of α -thalassaemia, are a result of either a single point mutation or a gene deletion of one of the globin *structural* genes. Therefore, hopes that control mutations would be common among the thalassaemias do not appear to be fulfilled. Even cases of 'hereditary persistence of fetal haemoglobin', reported by D. Weatherall (University of Oxford) and S. Ottolenghi (University of Milan), show linkage to the β -globin gene, and therefore are probably caused by a similar type of mutation.

The one new application of molecular biology to antenatal diagnosis was reported by Y.W. Kan (and originated from John Wilson, University of Georgia). He noted that the point mutation causing sickle cell disease, in the codon specifying amino acid 6 of the β -globin chain, eliminates a restriction enzyme site for *DdeI*. Wilson has used this in a direct analysis of DNA from patients and from the amniotic fluid of a fetus at risk. Previous DNA antenatal diagnoses have depended upon linkage between a restriction site polymorphism and the gene defect, and therefore have required a family study (Kan and Dozy *Proc. natn. Acad. Sci. U.S.A.* 75; 5631, 1978; Little *et al. Nature* 285; 114, 1980). If the defect itself can be examined directly at the nucleotide level, it simplifies matters

enormously for the specialist in molecular medicine, as both the theoretical problems of gene cross-over and the practical difficulties involved in family studies and the establishment of paternity all vanish.

Perhaps the most interesting view of the meeting concerned the attempts of clinicians in Italy and Greece to implement community programmes to control the disease. Pride of place must surely go to A. Cao (University of Cagliari), the organiser of the meeting, whose efforts in southern Sardinia have cut the incidence of thalassaemia from 1 in 250 live births to 1 in 800. This dramatic threefold reduction is not due exclusively to prenatal diagnosis, but it was very illuminating to learn that even in Sardinia, as a result of intensive popular education programmes, termination of pregnancy where prenatal diagnosis shows the fetus to be affected by this very severe disease is accepted in 95 per cent of cases. This surely indicates that community attitudes to very serious diseases are changing even where one might think that religious attitudes would prevent a successful programme.

D. Loukopoulou (University of Athens) is attempting a similar programme in Greece, but the fact that the country (and the problem) is so much bigger, and the absence of a unified health service and education programme, have meant that his programme has had less impact. However, he is now carrying out over 500 antenatal diagnoses a year, and in some regions in Greece the incidence of the disease is falling. He commented that the fact that affected children are now receiving blood transfusions (up to 20 a year) and very expensive support therapy to remove the iron from their bodies means that there is enormous financial and clinical burden on Greek paediatric departments which did not exist when most children died in the first two years of life. Therefore, the very fact that clinical care has improved (although still quite stressful for the patients) has created a dramatic need for better antenatal diagnosis.

B. Modell (University College Hospital, London) then reported upon the reaction from the Cypriot community in London to

Bob Williamson is Professor of Biochemistry at St Mary's Hospital Medical School, University of London.

the availability of antenatal diagnosis. There has not been a single case of β -thalassaemia within the Cypriot community during the past two years, as a result of intensive community education programmes and the awareness of both the community itself and those general practitioners in North London who serve it.

Parents who were at risk of having an affected child had virtually stopped reproducing before antenatal diagnosis, but now are having a normal number of children, as it can be guaranteed that all of these will be healthy. In this way, it has been shown that far from resulting in a decrease in the number of normal children

being born, antenatal diagnosis with abortion of affected pregnancies has resulted in a large increase in normal children, which is what both the families at risk and the community wish.

However, antenatal diagnosis (whether using fetal blood or DNA) can only deal with the situation in advanced countries at the moment, and even in Britain, Italy and Greece there are still seriously ill children with thalassaemia. Bone marrow transplantation for blood diseases such as the leukaemias is now an established, if risky, clinical practice in several centres, and D. Thomas (University of Seattle) discussed whether this procedure might also work for thalassaemia. The best results in bone

marrow transplantation are achieved with HLA-compatible brothers or sisters, and these certainly exist in thalassaemic families. However, transplantation should be carried out early, when the child is well, and not when the tissues have already been damaged by high levels of transfusion and the iron deposits this leads to. The question facing clinicians in the immediate future is whether to risk the life of a relatively well child with thalassaemia by giving a marrow transplant, knowing that if it works (a 75 per cent chance) the child will be healthy permanently — or to hope for other, less heroic treatment to be developed, knowing that if you guess wrong, the child may die in late teens of a very unpleasant disease. □

Minor body mass determination

from David W. Hughes

SPACECRAFT missions to both asteroids and comets have been much in the news recently. One of the aims of these investigations will be the accurate determination of the masses of minor bodies in the Solar System.

From theoretical considerations, the mass of a typical short-period comet has been estimated to be about 10^{16} g. It seems that the direct determination of a comet's mass by observing the perturbation that it induces in the orbit of another celestial body is quite out of the question. One of the closest approaches to Earth of any recorded comet was that of periodic comet Lexell, which on 1 July 1770 came to within 2.25×10^6 km, about six times further away than the Moon. Laplace recorded in his *Traité De Mécanique Céleste* (Paris, 1805) that this close encounter had no measurable effect on the length of the sidereal year and thus the mass of this comet must be less than 2×10^{-4} the mass of the Earth (which is 6×10^{27} g).

Asteroids are generally bigger than comets and therefore the problem is slightly easier. A great many positional observations are available over a long time period for the largest two, Ceres and Pallas, beginning with the years of discovery, 1801 and 1802. These observations provide information about the accumulated gravitational effects which each of the two bodies have produced in the orbital motions of the other. The total effects are small. Schubart (*Astr. Astrophys.* 30; 289, 1974) integrated the data for 1802–1970 and found the mass of Ceres to be $(1.17 \pm 0.06) \times 10^{24}$ g. Pallas was smaller — $(2.6 \pm 0.8) \times 10^{23}$ g — and so the error in its mass determination was relatively larger.

A more recent determination of the mass of a minor body has been attempted by analysing the radio-tracking data of a spacecraft as it flies close to the body. The Viking Orbiter 1 made several flybys to

within 80–300 km of the Mars satellite Phobos. Radiometric tracking from Earth measured the spacecraft – Earth distance and the geocentric velocity of the spacecraft. The latter was achieved by recording the Doppler shift in the S and X band frequencies transmitted by the spacecraft. Analysis of the gravitational perturbations induced in the orbit of the spacecraft led to a mass determination of $(9.9 \pm 0.6) \times 10^{18}$ g for Phobos — an object which is about 21.5 km across.

The problem of measuring the mass of a comet by using the tracking data of an *in situ* spacecraft has been analysed by D.K. Yeomans, M. Ananda, W.L. Sjogren and L.J. Wood of the Jet Propulsion Laboratory, Pasadena (see *J. Astronaut. Sci.* 29; 19, 1981). In a representative experiment, the authors decide to determine the mass of the short-period comet Tempel 2 during a rendezvous mission. A solar electron propulsion system (SEPS) has been used to transfer the spacecraft to the comet and the data are collected near perihelion when the comet becomes active. A single slow flyby on the sunward side of the cometary nucleus is assumed.

The main relative spacecraft–comet acceleration is due to the gravitational attraction between the two bodies. There are, however, four additional accelerations to be considered. Comets near perihelion give off a lot of gas and dust, at a rate that varies with heliocentric distance and is probably uneven, the material coming from a series of relatively short-lived active regions on the spinning cometary nucleus. The magnitude of the acceleration effect can thus vary stochastically about some mean value. It also depends on the orientation of the solar array panels, obviously

being larger when these are face-on to the flux as opposed to being 'feathered'. Solar radiation pressure also influences the spacecraft and again the effect depends on the area of the spacecraft presented to the radiation flux. Radiation pressure is generally larger than the gas and dust drag but is more easily accounted for as it can be considered constant over short time periods. There is also a tidal acceleration due to the unequal distance between the spacecraft, the comet and the Sun. Finally, spurious accelerations would be present, generated by the spacecraft due to the incomplete shut down of SEPS engines, altitude and articulation control subsystem gas jets. The operation of tape recorders and resonant vibrations of large solar panels also cause problems.

Two primary measurements are taken. Doppler tracking using the Deep Space Network of tracking stations enables the velocity to be found to an accuracy of 0.1 cm s^{-1} over a count time of 1 min. Spacecraft–comet ranging using, for example, X band (10 GHz), a pulse length of $2 \mu\text{s}$ and a pulse repetition frequency of 250 Hz can give an accuracy of 20–30 m when the spacecraft is within 1,000 km of the cometary nucleus.

Yeomans and colleagues find that the Doppler tracking data are more useful than the spacecraft–comet ranging data when it comes to determining the mass of the comet. Accuracy is improved by having both. The sensitivity of the mass determination to the stochastic nature of the gas and dust emission indicates that an accurate on-board accelerometer would also improve the accuracy.

Returning to Tempel 2, if the rendezvous spacecraft flies by the comet at a minimum distance of 100 km at a velocity of 2 m s^{-1} and data are collected for about 8 days from both a Doppler and a ranging system, a conservative estimate is that the mass can be measured to just better than 15%. □

David W. Hughes is in the Department of Physics, University of Sheffield.

Leeuwenhoek's specimens discovered after 307 years

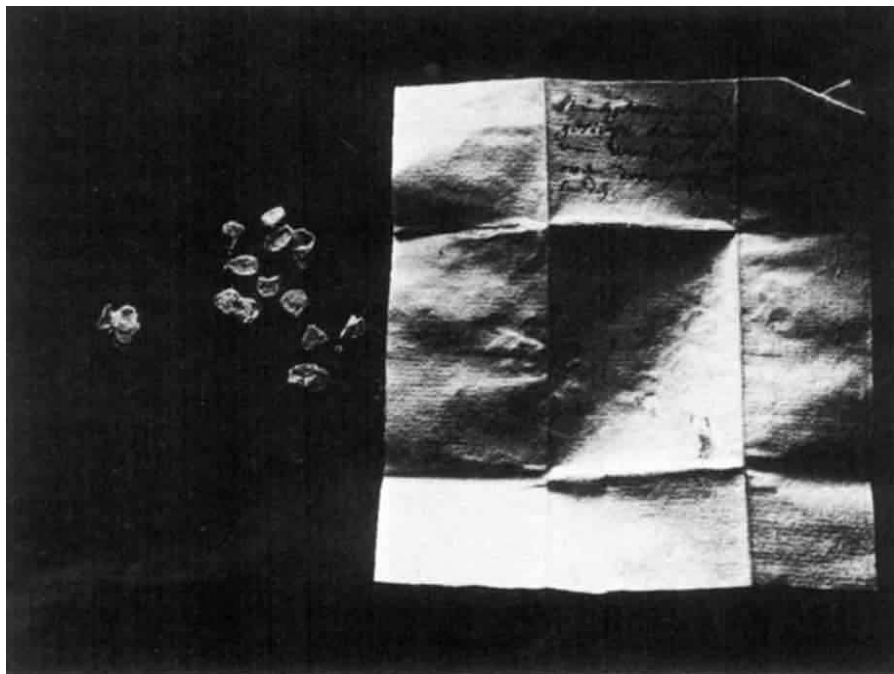
from Brian J. Ford

AFTER not being seen since 1674, original specimens sent by the "father of the microscope", Antony van Leeuwenhoek, have been found to be still in existence. A systematic search through the four large files of his letters sent to the Royal Society, revealed nine little envelopes containing original material and fixed to the final pages of three of the letters. Only minute traces of one of the specimens ("t Witte van een schrijff-penne" — white from a writing-pen [quill]) remained, but the remainder were intact. The sectioned material was compressed into compact masses, but otherwise the specimens were in excellent condition.

Antony van Leeuwenhoek (1632–1723) lived and worked in Delft where he produced his own single-lens microscopes. His status as the founder of microscopy derives from decades of dedicated and accurate observation, faithfully recorded in lengthy letters most of which were sent to London. His observations of spermatozoa, bacteria and a host of histological specimens were without precedent. Many workers have since commented on the 'crudity' of his techniques, emphasizing that his material was examined whole, or 'torn apart', and it has become generally accepted that section-cutting did not begin until the middle years of the nineteenth century.

The discovery of his specimens reveals two examples of plant material cut as fine sections. They are "Pit van vlier" — elder pith, and "Kurk" — cork. The conclusions that can be drawn from these specimens about seventeenth-century microtechnique are many (a fuller account by the author appears in the current *Notes and Records of the Royal Society*) but the central finding is that van Leeuwenhoek, working at the dawn of microscopy, could prepare by hand sections that would be acceptable for laboratory use today. □

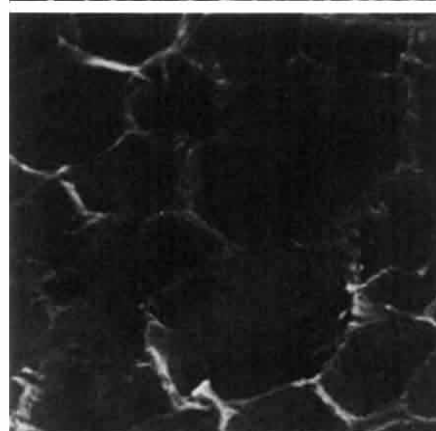
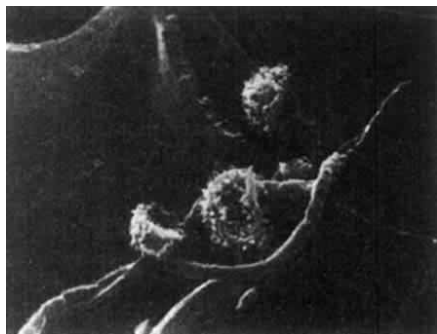
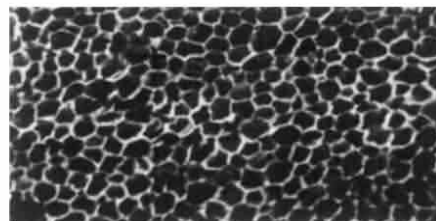
Brian J. Ford is in the Science Unit, Cardiff.



"Stuckjens vande gesicht senuwe van een koebeest over dwars afgesneden" — pieces of optic nerve of a cow cut in transverse section. The specimens were attached to a letter dated 1 June 1674, and signed by van Leeuwenhoek, as shown below.



Right, the stereoscan microscope reveals the regular, fine cut of the sections of cork. Though the hand-cutting technique seems primitive, the results are good enough for present-day microscopy. Below right, the sections of cork shown reveal the room-like appearance from which the term 'cell' was derived in 1663.



Left, high-magnification examination of the cell walls reveals occasional fungus spores. Thread-like hyphae occur from time to time in the sections, but have caused no structural change over 307 years. Right, specimens of papery material which van Leeuwenhoek claims arose from water are attached to a letter dated 17 October 1687. Algal cells and diatom chains revealed in this scanning electron micrograph confirm his finding.

Mixed-species foraging groups

from Jared M. Diamond

A bird-watcher is walking through a tropical rainforest, dejected at having seen and heard almost no birds for the last hour. Then a faint sound becomes audible afar, gradually grows louder and approaches. Soon the forest comes alive with dozens of bird species busily foraging together and calling simultaneously. The feverish excitement lasts for a few minutes, dies down as the flock moves on, and the forest is again silent and seemingly devoid of birds. A mixed-species foraging party has passed.

Analogous scenes repeat themselves daily on all the continents and in all the oceans, in the treetops, on the ground and underwater. The most familiar mixed foraging parties involve birds; the most spectacular ones, African open-country grazing mammals. Further examples are being discovered among fish, cetaceans and primates, and in at least six cases birds and primates have been found regularly associated with each other. The growing literature on mixed foraging parties (flocks, herds, or schools) has focused on four main questions. Are flocks chance aggregations or always composed of the same individuals? Who leads, and who follows? Why do different species associate? Why do the species sometimes resemble each other in appearance and call far more closely than one would expect from their taxonomic relationship?

Recent papers on mixed bird flocks of the neotropics¹⁻⁴ have revealed how unexpectedly intimately the lives of flock members belonging to different bird species are intertwined. These authors introduced to the study of mixed flocks techniques for recognizing individual birds by capturing, colour-banding and releasing them. Using these methods they found that the core of a flock consists of individuals of different species that forage together daily for many years, probably for their whole adult lives. Up to a dozen core species are each represented by one pair or family, whose territorial boundaries coincide among the core species and which may roost at night in the same tree. Even territorial confrontations are shared: when one flock happens to meet the neighbouring flock at the territory boundary, the Bluish-slate Antshrikes, Dusky-throated Antshrikes, Dot-winged Antwrens and White-flanked Antwrens of one flock simultaneously display and call to their conspecifics in the other flock. Individuals of additional flock species with smaller territories than the core members join the flock only when it enters their territory. Species with larger territories than the core members forage first with one flock, then with a neighbouring one a few hours later. Some species switch between

two distinct types of flock that occur in the same area: a slowly moving flock in the understory basically consisting of insectivorous antbirds, and a rapidly moving flock in the canopy of omnivorous tanagers and honeycreepers.

Are some species leaders and others followers? Usually there is a leader species (the 'nuclear species'), which is conspicuous in its plumage, frequent vocalizations and nervous behaviour such as wing- or tail-flicking, and which forages in groups of several individuals⁵. Nuclear species of bird flocks are often tits in Europe, North America and west Africa, babblers in tropical Asia, thornbills and fairy-wrens in Australia, gerygone warblers in New Guinea, and antbirds and tanagers in the neotropics (see, for example, ref. 6). Other species follow the conspicuous noisy whirlwind of the nuclear species group. In most bird-mammal associations, the mammal is the leader, the bird the follower, but tamarine monkeys of the neotropics are thought to follow tyrannid flycatchers.

Flocking seems to serve several purposes simultaneously, and different motives may predominate in different cases. Some of the leading theories are the convoy theory, gang theory, beater theory, pirate theory and feeding efficiency theory.

Convoy theory. Flocks may be an anti-predator device, like the World War ship convoys that were so effective against U-boats^{7,8}. The more individuals in a flock, the more eyes and ears there are to detect a predator, the lower is the chance of a given individual being the victim of an attack, the harder it is for the predator to focus on any target amidst all the activity and the better is the flock able to mob or ward off the predator. Observations that support the convoy theory include that lone birds peer about more, and spend less time feeding, than birds in a flock⁹. On African plains, near-sighted zebras with acute hearing associate with far-sighted species such as ostriches, wildebeest and giraffes, so that each complements the other's detector capabilities. Near-sighted gleaning birds tolerate the proximity of far-sighted salliers that occasionally rob them of food (antwrens tolerating antshrikes in the neotropics, tits tolerating drongos in Africa): the improved predator detection is worth the lost food.

Gang theory. Flocking may permit flock members access to resources within a territory whose owner would be able to expel the members if they were solitary. Thus a flock is like a gang of marauders overwhelming a victim by their numbers.

Jared M. Diamond is Professor of Physiology at University of California Medical School, Los Angeles.

For example, on Caribbean reefs the Three-spot Damselfish can inhibit single Stripe Parrotfish from feeding in its territory, but not parrotfish schools¹⁰. Similar cases are reported for Australian honeyeaters, Australian corvids and North American mockingbirds^{11,12}.

Beater theory. Tiger hunters follow a line of beaters who create a disturbance and flush out tigers. Similarly, some flock members, in the course of their own foraging, flush prey that a different species can capture. For example, gleaning and bark-feeding birds disturb resting insects into flight, where a sallying bird may catch them. Drongos, which are sallying insectivores, are notorious for following large or active animals so that they can catch insects disturbed by these beaters. Beaters used by drongos include elephants, giraffes, buffalo, ostrich, baboons and occasionally humans in Africa; monkeys in Borneo and Celebes; and babblers in New Guinea. In the neotropics, monkeys are used as beaters by the hawk *Harpagus bidentatus*. Africa's Ground Hornbills follow baboons that are turning over logs. African arboreal monkeys that eat fruit husks are joined by squirrels that eat the fruit kernels, ungulates that eat the fruit falling to the ground and hornbills catching insects flushed out by the monkeys.

Pirate theory. From using other species as beaters to flush out food that the beaters could not take, it is a small step to the role of pirate: seizing food that another flock member has caught. Drongos practice occasional piracy on other birds in New Guinea and African flocks¹³; antshrikes victimize antwrens in neotropical flocks. Why does the victim tolerate the pirate? Probably because the pirate is a noisy, far-sighted sallier that detects and warns of predators.

Feeding efficiency theory. Species that flock together have similar diets, although they differ in foraging technique, height and substrate^{9,14,15}. At first it seems in defiance of competition theory that species with similar diets should flock, and that flocking should vary inversely with food density^{16,17}. Yet flocking promotes feeding efficiency in several ways. The more individuals searching, the more likely is a good feeding patch to be found. Animals can learn new foraging techniques from conspecifics¹⁸: Morse saw a Blue Tit fly up to a Lesser Spotted Woodpecker pecking at bark and begin pecking next to it. Perhaps more important, by travelling together, species with similar diets can keep track of where food has already been gathered and is no longer worth seeking¹⁷. Imagine five janitors who clean a large hall and are paid per kilogramme of trash collected. One janitor uses his hands, one a pitchfork, one

a broom, one a vacuum-cleaner and one a carpet-sweeper. Each janitor is better adapted than the others to collecting certain trash items, but there are many items that can be collected by any of several janitors. A good strategy for all janitors is to forage together, thereby constantly assuring themselves of trash-rich areas and avoiding areas partly cleared by other janitors.

For whatever reason species do forage together, it is striking that flock members of widely different families often resemble each other in appearance, vocalizations, or both. As examples of this co-evolutionary phenomenon, termed social mimicry, Moynihan¹⁹ noted that flocking tanagers in the mountains of Panama are mainly black and/or yellow, in the northern Andes brilliant blue or blue and yellow, and in the south-central Andes blue-grey dorsally and brown ventrally. A prime example of this tendency for birds of a feather to flock together is the 'brown and black flock' of New Guinea, first noticed by the flamboyant Italian explorer Count Luigi d'Albertis in 1880 and rediscovered much later by Bell²⁰. The flock involves up to 22 medium-sized, omnivorous bird species from seven families (birds of paradise, babblers, whistlers, drongos, cuckoo-shrikes, honey-eaters and flycatchers). All these species are either mainly rufous or mainly black; in three, the male is black, the female rufous. The flock is led by a group of the babbler *Garritornis isidori*, of which one individual is the actual leader and has a distinctive 'leader' call while his/her conspecifics have a 'follower' call. At least five other species in the flock mimic one or other of the babblers' calls.

Why should species from such different families converge in plumage? We are not dealing here with detailed mimicry of plumage patterns, but merely broad resemblances in overall colour. Moynihan points out that interspecific signals are required so that mixed flocks can form and maintain cohesion. The convergence in these signals may be for economy: it is much easier for a flock member to stay with the flock and avoid wandering off with some non-flock species if it need only remember to look for brown or black birds, instead of keeping 21 different colour patterns in mind. □

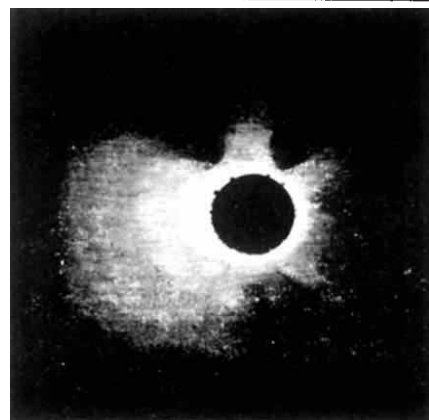
1. Gradwohl & Greenberg *Auk* 97, 385 (1980).
2. Munn & Terborgh *Condor* 81, 338 (1979).
3. Powell *Auk* 96, 375 (1979).
4. Buskirk *et al.* *Auk* 89, 612 (1972).
5. Moynihan *Smithson misc. Collns* 143, 1 (1962).
6. Bell *Emu* 80, 227 (1980).
7. Hamilton *J. theor. Biol.* 21, 295 (1971).
8. Pulliam *J. theor. Biol.* 38, 419 (1973).
9. Morse *Ibis* 120, 298 (1978).
10. Robertson *et al.* *Ecology* 57, 1208 (1976).
11. Paton *Emu* 80, 213 (1980).
12. Merritt *Auk* 97, 869 (1980).
13. Greig-Smith *Ibis* 120, 284 (1978).
14. Wiley *Auk* 88, 881 (1971).
15. Austin & Smith *Condor* 74, 17 (1972).
16. Morse *Ecol. Monogr.* 40, 119 (1970).
17. Cody *Theor. Populat. Biol.* 2, 141 (1971).
18. Rubenstein *et al.* *Ibis* 119, 10 (1977).
19. Moynihan *Evolution* 22, 315 (1968); *Am. Nat.* 177, 372 (1981).
20. Bell *Emu* 76, 95 (1967).



THE CHEMISTRY OF THE SUN

Although in the year 1866 a great many people were familiar with the spots on the sun, those who had been favoured by a sight of a total eclipse, and many more who had read the accounts of total eclipses, knew that there was a great deal more of the sun than one generally sees. From the time of Stannyan, who observed the prominences at Berne, down to the year 1842, let us say, several eclipses had been observed, and very beautiful coloured phenomena had been recorded by different observers. Red things had been seen projecting round the dark moon during the time of eclipse, and although many held them to be beautiful effects produced by the passage of the moon over the sun, or even clouds in the atmosphere of the moon coloured by the strange way in which the solar light then fell upon them, a larger number of people, on the other hand, insisted that these things must really belong to the sun. Now if that were so, it was perfectly clear that we should not be contented with merely observing the chemical nature of the spots.

We have a round dark moon, which in this case is represented as entirely covering the sun; then these different prominences and luminosities, this wonderful set of streamers, or whatever you like to call them, which seem to veil, or to render less distinct, something else which is lying beyond them. You will see here that some of these prominences are red, and others have a yellow tinge, and that, quite independent of the colour of the prominences, we have the most exquisite coloured effects. Sometimes the radial structure is not so marked, and reveals



Eclipse of 1870. Photograph of the corona taken at Syracuse.

indications of structure further away from the sun. You see wonderfully delicate tracery, lines being seen now in one part and now in another. In the photograph taken during the eclipse of 1870 we see that the luminosity of the solar atmosphere was excessively irregular, by which I mean that in one part we get a very considerable excess of light, quite independent of the sharply defined prominences, whereas in other portions the atmosphere of the sun at the same height is not nearly so luminous. Now in none of these cases have we been able to see the thing which struck us most clearly the moment the artificial eclipse system was set at work. It is a good indication of the extreme difficulty of making observations during eclipses, and how important it is that one should have a method which makes us independent of them.

From *Nature* 24, 21 July, 271, 1881.

Origin of the galaxies

from Joseph Silk

COSMOLOGISTS generally believe that galaxies formed as a consequence of the growth of tiny density fluctuations in the very early Universe. The remarkable uniformity of the cosmic microwave background radiation has provided convincing evidence for the high degree of regularity of the Universe at an early epoch but within the past year two groups of astronomers, at Florence and Princeton, have reported evidence for a very weak anisotropy on an angular scale of 90°, which is probably produced by similar matter fluctuations to those from which the galaxies evolved.

One of the most tantalizing issues in cosmology concerns the origin of these fluctuations. It is thought that the initial conditions of the big bang might provide an answer. Perhaps the Universe is the way it is because of the way it was. Surely the resourceful cosmologist can improve on this hollow echo of creationism, for if the initial conditions were very different, then we would not be here to take stock of the situation. This concept of the observer's role in determining the state of the Universe has been elevated by Robert Dicke and Brandon Carter to the status of a fundamental cosmological principle: the anthropic principle. Its power arises

because the growth of fluctuations is inevitable from the beginning of time.

Consider the history of a fluctuation destined eventually to form a galaxy. It grew in amplitude throughout much of the early expansion of the Universe. The earliest epoch that the cosmologist can usefully discuss is the Planck instant, a mere 10^{-43} second after the singularity. At this time, our galaxy-sized fluctuation must have had an infinitesimal but non-zero density contrast. These inhomogeneities are best considered as fluctuations in the spatial curvature, or slight wrinkles in the geometry of space-time. The magnitude of such primordial wrinkles must be about 1 part in 10^3 , if fluctuations in the energy density are considered; for fluctuations in the matter component alone, the associated wrinkles attain a considerably smaller but well defined value. Were the fluctuation amplitudes much smaller, galaxies would not have formed by now; were they much larger, eventual collapse in the early Universe would have formed separate topologically disjoint universes, bearing little relation to our observed universe.

Although this argument purports to show that the initial conditions may have been unique, it certainly does not provide any explanation. One is left with the following conundrum. The very early

Joseph Silk is in the Astronomy Department at University of California, Berkeley.

Universe was uniform to a high degree. However to form galaxies, fluctuations must have been present at an exceedingly small but non-zero level. This situation is aesthetically unappealing — granted that the early Universe must have been very nearly uniform, one would obviously prefer to begin with a perfectly homogeneous and isotropic expansion, and see fluctuations develop spontaneously at some later epoch. No one has been able to discover a means of generating fluctuations over a sufficiently large scale to form galaxies in the standard big bang cosmology. To some cosmologists, this failure has provided sufficient reason to reject the standard model. Some have opted for an initially cold big bang in the hope that cold matter may be more susceptible to fragmentation. Others have taken a stronger stance, and rejected the big bang entirely.

A fascinating resolution of this cosmological dilemma has been recently proposed by Ya. B. Zel'dovich of the Space Research Institute, Moscow (*Mon. Not. R. astr. Soc.* **192**; 663, 1980) and Alexander Vilenkin of Tufts University (*Phys. Rev. Lett.* **46**; 1169, 1981). Recent applications of elementary particle theory to the very early Universe imply the possibility of a novel mechanism for the spontaneous generation of the seed fluctuations from which galaxies eventually formed. The origin and nature of elementary particles can be understood in terms of a class of theories known as the grand unified theories (GUTs), which seek to unify the electromagnetic, weak and strong nuclear forces into a single theoretical framework. Above the grand unification energy of about 10^{15} GeV, these fundamental forces of nature all play an equal part, and matter is completely symmetrical in all its properties. The very early Universe provides a natural laboratory for testing GUTs, since particle energies attained then greatly exceed those produced in any man-made particle accelerator. One of the most elegant consequences of GUTs is that they predict (at least in order of magnitude) the observed baryon asymmetry of the Universe. The breakdown of symmetry occurs some 10^{-35} second after the big bang, when particle energies have dropped below the grand unification energy. This produces a slight excess of particles over antiparticles which is crucial to the existence of an environment in which life could evolve, for the fact that we and our surroundings consist exclusively of matter rather than antimatter implies that in the very early Universe, when the cosmic blackbody radiation copiously produced particle-antiparticle pairs, the amount of matter exceeded that of antimatter by about 1 part in 10^9 . Our very existence is therefore a consequence of GUTs and the breakdown of symmetries as the Universe expanded and cooled, for a Universe which precisely conserved matter-antimatter symmetry would have undergone

considerable annihilation; little residual matter (or antimatter) would remain.

Zel'dovich and Vilenkin argue that GUTs may have profound implications for the large-scale structure of the Universe. The breakdown of symmetry at 10^{-35} second can be regarded as a phase transition. A useful analogy is an isotropic ferromagnet, which when cooled below a critical temperature (the Curie point) spontaneously develops magnetism. The direction of magnetization is indeterminate and random, and may vary in different regions or domains. The GUTs phase transition may involve the spontaneous generation of a domain structure. The symmetry breakdown is believed to be effected via an intermediary class of particle — the Higgs particle — and its associated fields. These are important in the unification of strong and electroweak interactions, being responsible for the masses of the elementary particles once the symmetrical state is broken. As the Universe cools, the symmetry breakdown occurs because particles and radiation are no longer energetic enough to create the full panoply of states of all types of particle predicted by GUTs. The end state of the matter is characterized by a vacuum state of lowest energy. In the classical physics regime, the vacuum has approximately zero energy, although small fluctuations are present. The fluctuations are variations in the energy of a reservoir of virtual pairs of particles and antiparticles that are released under extreme conditions of temperature or gravitational stress. The vacuum state after the GUTs phase transition has a finite energy density associated with the Higgs field. It is believed that the Universe began with a very large vacuum energy density of almost equal magnitude but opposite sign, forming a false vacuum state, so that one ends up after the phase transition in the true vacuum of zero energy.

In fact, the true vacuum state is not unique. Although the minimum energy state of the Higgs field has a universal value, implying the same energy density everywhere, the direction of the associated field is indeterminate. Multiple true vacuum states are possible, depending on the degree of complexity of the Higgs field. The phase transition can result in the formation of different domains of vacua, corresponding to the multiple vacuum states. The direction but not the magnitude of the Higgs vacuum expectation value differs in each domain. The domain boundaries are characterized by gradients in the Higgs field, and therefore contain energy density; they also consist of residual false vacuum that has not undergone the phase transition. When the Higgs particles have decayed, the Universe contains the remnants of these boundaries in the form of topological structures of false vacuum that may contain a very considerable energy density.

Theory indicates that possible forms for

the domain boundaries are topological discontinuities that can be either two-, one-, or zero-dimensional in character. There are two-dimensional walls, one-dimensional strings, or monopoles, which are point-like singularities. An important characteristic of the energy density associated with the monopole's structure is that it effectively behaves as a rest-mass, and assumes a progressively more important gravitational role as the radiation is redshifted away. The walls and strings suffer very little friction from the ambient radiation and matter, and the tendency of these initially highly convoluted structures to straighten out causes them to move at near the speed of light on the horizon scale. Density fluctuations are generated on and below the horizon scale, and larger and larger fluctuations develop as the horizon scale grows.

In fact, expanding walls can be excluded by simple astrophysical considerations. Zel'dovich, Kobzarev and Okun pointed out (*Zh. Eksp. Teor. Fiz.* **67**; 3, 1974) that such walls should be extremely massive, and would grossly perturb the cosmic microwave background radiation. However strings are quite another matter. Their one-dimensional structure implies that their influence in the present epoch is relatively modest. Strings may still lead to sizable density fluctuations. The process of generating density fluctuations continues after the phase transition at 10^{-35} second, perhaps until the present epoch if the strings can survive. At formation, the strings are very tangled, and both infinitely long strings and closed loops are present. Dissipation tends to reduce the tension in the strings on scales smaller than the horizon, and the strings gradually straighten out. New strings are likely to form as intersecting strings form closed loops; this provides a principal means of straightening the strings. The expansion of the Universe stretches the strings on scales larger than the horizon, and the characteristic dimension of a string at any epoch is in fact the horizon scale.

The density fluctuations generated by the strings can provide the seeds for galaxy formation. The mass of a galaxy is first contained within the horizon at an epoch of about 10 yr and the mass of a galaxy cluster after about 10^4 yr. Strings will be present on the horizon scale moving at light speed and their gravitational pull will induce density fluctuations. Zel'dovich argues that the density of infinite strings remaining from the GUTs phase transition amounts to about 0.1 per cent of the mean cosmological density. This implies that the expected amplitude of density fluctuations is of the order of 0.1 per cent. Such fluctuations survive the radiation even when their growth is inhibited by the frictional effect of radiative viscosity. After the decoupling epoch, the fluctuations grow in amplitude uninhibited by the radiation, accreting the surrounding matter. Galaxy formation eventually occurs once the fluctuations become of

very large amplitude and collapse.

How inevitable is such a scenario? It seems that in general types of GUT, these topological singularities (either walls, strings or monopoles) will form as the symmetry breakdown occurs, depending on the degree of complexity of the Higgs field. The only requirement is that the Higgs field possess a finite correlation length, beyond which its expectation value in different locations is completely uncorrelated. This provides our best hope for understanding galaxy formation in an initially homogeneous and isotropic universe. Galaxies may therefore have originated from a tangle of strings spontaneously produced some 10^{-35} second

after the big bang. The strings are best visualized as topological knots of high energy density that may be either of infinite length or closed loops. As the Universe expands, the strings untangle, continuously generating sizable density fluctuations via their gravitational effects on scales less than the instantaneous horizon scale.

Cosmologists are not entirely convinced by this scheme, since it involves rather specific assumptions about elementary particle theory concerning the nature of the Higgs field that have not received widespread acceptance. One difficulty is that not only strings are produced by the GUTs phase transition. Massive monopoles are

predicted by the simplest GUTs, and these particles could dominate the present mass density of the Universe, contrary to observation, unless the theory of the very early Universe is carefully adjusted. Perhaps strings will only meet universal acceptance if more direct proof of their existence is forthcoming. This is most likely to emerge from their effect on electromagnetic radiation, in the large-scale angular anisotropy of the cosmic microwave background radiation. Another intriguing possibility is that the gravitational lenses, recently identified with close pairs of quasars having strong spectral similarities, may be due to the gravitational deflection of light by cosmological strings. $\square$

Bose–Fermi equivalence and soliton theory in solid-state physics

from R. K. Bullough

NONLINEAR oscillators and nonlinear wave mechanisms are providing the key to our understanding of an increasing number of fundamental processes in physics, chemistry and biology. Some years ago nerve axon pulse propagation was modelled by the nonlinear Hodgkin–Huxley or FitzHugh–Nagumo equations. Now nonlinear wave equations are providing insights into, for example, the way chemical concentration waves may implement the changes in anatomical structure demanded by a mutant gene, or cell division in the emerging embryos of sea urchins.

Although most of these nonlinear systems can be handled by the theoretician only by use of large computing machines, there is a large and growing class of nonlinear wave equations which so far have been applied in theoretical physics rather than theoretical biology, and which can be solved analytically, that is, in terms of the familiar mathematical functions. Many of these systems have soliton solutions—the subject of this article.

A word first on nonlinearity; even at school we learn to write down the equation of motion for the angular displacement u of a simple pendulum in the form $Ml \, d^2u/dt^2 = -Mg \sin u$. Then we immediately linearize this equation by replacing $\sin u$ by u and so gain the harmonic oscillator equation $d^2u/dt^2 + (g/l)u = 0$ which has $\cos(g/l)^{1/2}t$ or $\sin(g/l)^{1/2}t$ as solution. But the linearized theory is applicable only if u is small. So what happens if the pendulum is rigid and can turn right over so that u is 2π or bigger? And what is the situation if we couple such oscillators together in a regular array to form a model crystal? In linear approximation we gain the so-called Klein–Gordon (KG) equation, $u_{xx} - u_{tt} = m^2u$, which is familiar (the

notation is the convenient one $u_{xx} \equiv \partial^2u/\partial x^2$, etc. and m here is a number playing the part of a mass). But the nonlinear analogue is the sine–Gordon (s–G) equation $u_{xx} - u_{tt} = m^2 \sin u$ (the name is no accident!). The s–G has soliton solutions but the KG does not.

This word ‘soliton’ was coined by Norman Zabusky and Martin Kruskal (*Phys. Rev. Lett.* **15**; 240, 1965) who were concerned with a very different-looking wave equation, the Korteweg–de Vries (KdV) equation, $u_t + 6uu_x + u_{xxx} = 0$, and in essence they observed that its solitary wave solution, $u = \frac{1}{2}V \operatorname{sech}^2\{\frac{1}{2}V^{1/2}(x - Vt)\}$ asymptotically preserved (or exchanged) both its form and its speed V in collision with other such solitary waves travelling at different speeds. Because of this particle-like property, these solitary waves were dubbed solitons. We now know that this behaviour typifies a whole class of nonlinear wave equations, outstanding examples being the KdV, the modified KdV (MKdV) $v_t + 6v^2v_x + v_{xxx} = 0$, the nonlinear Schrödinger equations (NLS) $i\psi_t = \psi_{xx} \pm 2\psi|\psi|^2$ (with ψ complex) and the s–G. Clearly the solutions of such equations could be superb examples of ‘elementary excitations’ in some context of many-body physics.

With this in mind, solid-state physicists, in particular, have been quick to use solitons to explain the high, anisotropic and essentially one-dimensional, non-ohmic electrical conductivity of certain organic polymers like TTF–TCNQ at temperatures $T \sim 3\text{K}$, or the central peak observed in the neutron-scattering spectrum from crystalline SrTiO_3 in the displacive regime, or the central peak in the spec-

trum of the effectively one-dimensional ferromagnetic crystal CsNiF_3 for $T \leq 10\text{K}$, and so on. Already in 1978 the conference report ‘*Solitons and Condensed Matter Physics*’ (eds A. R. Bishop and T. Schneider; Springer) was able to review a wide range of such applications.

The physical insight displayed in many of these investigations is admirable. Still it seems a pity that the word soliton is so often misused: sometimes it simply means a solitary wave (a pulsed solution with argument $(x - Vt)$) or even just a lump; jacobian elliptic functions like $\operatorname{cn}(x; k)$ (which tends to $\cos x$, $k \rightarrow 0$, or $\operatorname{sech} x$, $k \rightarrow 1$) are called ‘multisolitons’ or a soliton ‘lattice’. However, there is a mathematical structure surrounding the solitons and their collision properties which does not surround the simpler solitary waves; and this mathematical structure may be relevant to the physical problem in hand.

A case in point is the interesting paper entitled ‘*Soliton Lattice in Polyacetylene, Spin–Peierls Systems and Two-Dimensional Sine–Gordon Systems*’ by Baruch Horovitz (*Phys. Rev. Lett.* **46**; 742, 1981)—as well as a frequent misuse of the word soliton in this way in a series of papers by K. Maki and his collaborators who are referenced in the Horovitz paper. In his paper, Horovitz treats three physical systems: a fermion system concerned with the insulator-to-metal transition at $\sim 1\%$ doping observed in polyacetylene, which is explained in terms of a charge carrying soliton ‘lattice’ or soliton ‘drop-lets’ of that lattice; a ‘lock-in’ phase transition in a boson system engineered by quantum corrections to the classical s–G; and the first-order phase transition observed in the magnetic spin system TTF–BDT (Cu) [tetrathiafulvalene-bis-(1,2-perfluoromethylethylene-1,2-dithiolato)-copper] in a strong magnetic field ($\sim 120\text{kG}$) at low temperatures

R. K. Bullough is Professor of Mathematical Physics in the Department of Mathematics, The University of Manchester Institute of Science and Technology.

(~10K). The attraction of the analysis is that all three very different physical problems are shown to be mathematically equivalent. So it is a pity, and I think misleading to the physics, that the author does not make proper use of the mathematical structure surrounding the theory of solitons.

In his first example the author uses Peierls's idea (subsequently extended by Fröhlich as an incipient theory of superconductivity) that the electrons in a one-dimensional lattice with a half-filled band can lower their energy if a gap in the electron energy spectrum is induced by a suitable periodic distortion of the ion lattice. If the lattice distortion moves (as a phonon), the ions' charges stay locked to their sites; but if the lattice and its distortion have incommensurate periods, the electrons can travel freely with this phonon and constitute a resistanceless current. In practice this moving 'condensate' is pinned—typically by a periodic potential like $\cos \phi(x, t)$ depending on the condensate phase ϕ . In this particular case the current is carried by the soliton 'kink' (or 'antikink') solutions, $\phi = 4 \tan^{-1} [\pm \exp \{m(x - Vt)/(1 - V^2)^{1/2}\}]$ of the s-G. Such a theory describes the non-ohmic conductivity of TTF-TCNQ for $T < 4.2\text{K}$ quite well (see Cohen *et al. Phys. Rev. Lett.* **37**; 1500, 1976; and Trullinger *et al. Phys. Rev. Lett.* **40**; 206, 1978, and earlier papers). Note that the kink $\phi = 0$ as $x \rightarrow -\infty$, but jumps rapidly from 0 to 2π in a small region of order m^{-1} about $(x - Vt)$; the antikink jumps from 0 to -2π . Horovitz looks for a periodic lattice of kinks and antikinks in which the lattice ion displacement field $\Delta(x)$ swings (on some scaling) between $\pm 2\pi$ about the lattice sites. A sequence of s-G kinks and antikinks might describe just such a situation; however, such a multisoliton solution of the s-G at rest is not stable.

The mathematical problem in all three cases can be reduced to solving a Dirac-type eigenvalue problem for two-component spinor eigenfunctions $\psi_n(x)$ in the presence of the ion displacement field $\Delta(x)$ such that the total energy $E_0 = \sum \epsilon_n + \int dx \Delta^2(x)/2\pi\lambda$ is minimized; the $\sum \epsilon_n$ is the contribution of the eigenenergies of occupied electron states and λ a coupling constant. The eigenvalue problem takes the form $\hat{L}\psi_n(x) = \epsilon_n\psi_n(x)$ in which $\hat{L}$ is a Hermitean 2×2 matrix differential operator: $\hat{L} \equiv i\sigma_3 \partial/\partial x + \Delta(x)\sigma_1$ where σ_3 and σ_1 are Pauli matrices.

Experts in soliton theory will immediately recognize this eigenvalue problem as essentially that used by Zakharov and Shabat to solve the NLS in 1972. Likewise it solves the MKdV $v_t - 6v^2v_x + v_{xxx} = 0$. It looks uninteresting for solitons because the eigenvalues must be real ($\hat{L}$ is Hermitean) and we know that solitons are associated with complex eigenvalues. However, the point of all this here is the following: it is known that for the MKdV

(or for the NLS or the s-G for that matter) a 2×2 matrix $\hat{B}$ depending on v, v_x, v_{xx} and the ϵ_n can be found so that the MKdV equation is equivalent to the operator equation

$$\hat{L}_t \equiv [\hat{B}, \hat{L}] \equiv \hat{B}\hat{L} - \hat{L}\hat{B} \quad (1)$$

The pair of matrix operators $(\hat{B}, \hat{L})$ which makes equation (1) the MKdV is the familiar Lax pair for the MKdV—named after Peter Lax who first found this form for the KdV in 1968. To introduce it here we have extended $\Delta(x)$ by letting it depend on the parameter t and set $\Delta(x, t) \equiv v(x, t)$ so that $\Delta(x, t)$ follows the MKdV 'flow'. It is easily checked that equation (1) is the MKdV flow if, and only if, the ϵ_n s do not depend on t even though the scattering potential $\Delta(x, t)$ in $\hat{L}$ and the eigenfunctions ψ_n now do so. The MKdV is thus an isospectral flow in this sense. If a nonlinear equation like the MKdV can be put in the form of equation (1), we can expect to solve it for appropriate boundary conditions by the so-called inverse scattering method. This method generalizes the Fourier transform for linear problems to the present class of nonlinear problems via the isospectral transform $\hat{L}\psi_n = \epsilon_n\psi_n$. If the equation has soliton solutions we can find them through the complex eigenvalues ϵ_n in this way.

The key point is that the ϵ_n s remain stationary under the flow. Thus if $\Delta(x + t)$, of single argument $x + t$, is a particular solution of the MKdV for any t , we might guess that it will minimize the contribution of the electron eigenenergies ϵ_n to the total energy E_0 for all t ; furthermore, we know that under the MKdV flow with appropriate boundary conditions, $\int v^2 dx \equiv \int \Delta^2 dx$ is also a constant of the motion. Thus we guess that $\Delta(x + t)$, or scaling t , $\Delta(x + \gamma t)$ minimizes E_0 for an appropriate choice of γ .

Horovitz took a different route: he converts the spinor eigenvalue problem to a Schrodinger problem with scattering potential $\Delta^2 - \Delta'(x)$ where $\Delta'(x) \equiv d\Delta/dx$. He argues, incorrectly in my view, that this is of KG type and uses results in perturbation theory taken about the s-G equation by linearizing this equation about its kink solution to set $\Delta^2(x) - \Delta'(x) = U''(\phi(x))$ and to infer that $U(\phi(x)) \equiv \Delta_1^2(1 - \cos(\phi(x)))$. This potential is the s-G potential in its Hamiltonian

$$H = \int \left\{ \frac{1}{2} \phi_t^2 + \frac{1}{2} \phi_x^2 + m^2(1 - \cos(\phi(x))) \right\} dx \quad (2)$$

so Δ_1 is a mass m (Hamilton's equations for a field $\phi(x, t)$ generate the s-G from equation (2)).

Those familiar with the history of soliton theory will, however, see in the transformation $\Delta^2(x, t) - \Delta(x, t)_x = u(x, t)$ the celebrated Miura transformation which moves a solution $\Delta(x, t)$ of the MKdV to a solution $u(x, t)$ of the KdV $u_t - 6uu_x + u_{xxx} = 0$! It is a curious fact, associated

more with the theory of jacobian elliptic functions than with solitons, that the KdV has the solution $u(x + \Delta_1^2 t) \equiv u(\xi)$ given by $u(\xi) = \Delta_1^2 \cos \phi(\xi)$ providing $\phi'(\xi) = 2\Delta_1 \sin \frac{1}{2}\phi(\xi)$ so that $\phi'' = \Delta_1^2 \sin \phi$. Certainly this is the equation of motion (in the 'time' variable $\xi = x + \Delta_1^2 t$) of a pendulum capable of large angular displacements; and it has the static kink solutions $\phi = -\pi + 4 \tan^{-1} \exp \Delta_1 \xi$ in which the pendulum turns from upside down very slowly through angle 2π back to upside down. Nonetheless, it derives from a wave equation, the MKdV, very different from the s-G, and the physical associations it conveys are also very different.

Of course, like a good physicist, Horovitz still finds a correct mathematical solution for his lattice which for large period does indeed behave like a series of alternating kinks and antikinks of the s-G. He then finds that this lattice is stable for large doping levels (large values of the phonon-electron coupling constant λ) but becomes unstable for small λ . The formation of droplets of this lattice is suggested to herald the observed insulator-to-metal transition observed in polyacetylene.

In a similar way his so-called bose problem appeals to the fact that if the s-G hamiltonian equation (2) is linearized about a kink and the resulting small oscillations quantized as 'phonons', this quantum correction is essentially the energy of the fermion soliton lattice with a change in sign. Thus there is a first-order phase transition as the coupling constant λ is reduced. Similar considerations apply to his spin problem where $\pm \frac{1}{2}$ spins are carried by the solitons.

The moral of this interesting paper seems to be that good physical insight leads to good physics. But soliton theory is now so important in many areas of nonlinear theoretical physics that a good understanding of it has become an essential technical prerequisite for the practising theoretician. This is particularly true in high-energy physics where mathematical structure may be the only guide. It is a remarkable fact that the so-called spin $-\frac{1}{2}$ xyz lattice model can be mapped exactly, in the continuum limit of zero lattice spacing, onto the massive Thirring model (a relativistically co-variant fermion field theory) and that this maps exactly onto the quantized s-G model (a co-variant boson theory). This formal fermion-boson equivalence, as well as the remarkable mass spectrum of the quantized s-G, has stimulated much work on nonlinear field theories, especially the gauge and supersymmetric gauge theories, in high-energy physics. It has also recently stimulated a quantum version of the inverse scattering method for solving some of these quantized problems exactly. Apparently it was these ideas on fermion-boson equivalence that motivated Horovitz's line of analysis in his interesting paper on solid-state physics. □

REVIEW ARTICLE

Actions of cholera toxin and the prevention and treatment of cholera

Jan Holmgren

Institute of Medical Microbiology, University of Göteborg, S-413 46 Göteborg, Sweden

*The drastic intestinal secretion of fluid and electrolytes that is characteristic of cholera is the result of reasonably well understood cellular and biochemical actions of the toxin secreted by *Vibrio cholerae*. Based on this understanding it is possible to devise new techniques for the treatment and prophylaxis of cholera to complement those based on fluid replacement therapy and sanitation.*

CHOLERA patients have characteristically watery diarrhoea which leads to dehydration and metabolic acidosis. If untreated, this fluid loss rapidly leads to death. The disease is caused by an intestinal infection with *Vibrio cholerae*. These bacteria adhere to and colonize the small intestine and secrete an exotoxin—cholera toxin—that binds to receptors on the mucosal cells and stimulates intestinal adenylate cyclase activity. The resulting increase in cyclic AMP then causes diarrhoea and fluid loss by inhibiting uptake of sodium chloride by the villi as well as by stimulating active chloride secretion by crypt cells¹.

Koch, who identified *V. cholerae* as the causative agent of cholera, had in 1887 already proposed that the disease was toxin-mediated but it was not until 1959 that the Indian scientists De² and Dutta³ convincingly demonstrated the existence of a cholera toxin. This was purified^{4,5}, and its effect on the adenylate cyclase-cyclic AMP system soon established⁶⁻¹⁰. Since then, activation of adenylate cyclase by cholera toxin has been shown to occur in most mammalian cell types, the structure-function relationship of the toxin has been defined, the cell membrane receptor identified and, most recently, the mode of action of the toxin on adenylate cyclase explained in considerable detail.

This knowledge of the cholera toxin has made it a useful tool for cell biologists, biochemists and physiologists interested in aspects of cell membrane and cyclic nucleotide research that are essentially unrelated to cholera. It has also suggested various possibilities for novel approaches to prevention and treatment of disease. My main intention here is to describe the cellular action of cholera toxin and to discuss the possible exploitation of this knowledge in the design of vaccines, receptor-prophylactic agents, and antisecretory drugs against cholera.

The cholera toxin molecule

Cholera toxin is a protein with two types of subunit: a single 'heavy' subunit of molecular weight (MW) 28,000 noncovalently attached to a 58,000-MW aggregate of 'light' subunits (Fig. 1).

The demonstration that choleragenoid, a protein immunologically related to cholera toxin and able to bind to intestinal epithelium without having toxic activity⁴, contains the same 'light' subunits as the toxin but lacks the 'heavy' subunit strongly suggested that the 'light' subunits are responsible for cell binding (B subunits) and the 'heavy' subunit for the direct toxic activity (A subunit)¹¹⁻¹⁴. Experiments with toxin subunit fractions, prepared by gel filtration in acidic buffer and dialysed to allow renaturation and reassociation of subunits, confirmed that B subunit bound strongly both to the cell and to the isolated cell receptors, but was nontoxic. Purified A subunit neither bound to nor was toxic for intact cells. However, A reassociated with B had both binding and toxic activity in various whole-cell systems^{11-13,15,16}.

The requirement for membrane binding by the B subunits can be circumvented by using disrupted cells; in these conditions purified A subunit and also its A₁ fragment (see Fig. 1) can activate adenylate cyclase¹⁷⁻¹⁹. Furthermore, the characteristic lag period of 10–60 min observed in intact cells before any effect of cholera toxin on adenylate cyclase is seen⁶⁻¹⁰, essentially disappears in broken cells. Reduction of the disulphide bond between the A₁ and A₂ regions, but not the physical separation of the two fragments, seems to be necessary for the activity of A on adenylate cyclase²⁰.

The cholera toxin receptor

The first event in the action of cholera toxin on cells is the rapid, tight binding to receptors on the cell surface. Studies with ¹²⁵I-labelled toxin have shown that binding occurs almost instantaneously, is saturable and initially reversible^{15,21}. The number of binding sites per cell varies widely with the cell type but affinity of binding varies very little ($K_A \sim 1 \times 10^9 \text{ mol}^{-1}$) indicating that the receptor is the same for various cell types²¹⁻²³.

It is now known that the membrane receptor for cholera toxin is a specific ganglioside (Fig. 2). Van Heyningen *et al.*²⁴ observed that a crude ganglioside mixture inactivated cholera toxin; J.H. *et al.*²⁵, Cuatrecasas²¹ and King and van Heyningen²⁶ showed that this inactivation resulted from specific binding between the toxin and a single ganglioside, G_{M1}. G_{M1} neutralized cholera toxin in about equimolar proportions and gave a specific precipitation band with cholera toxin in gel-diffusion tests²⁵.

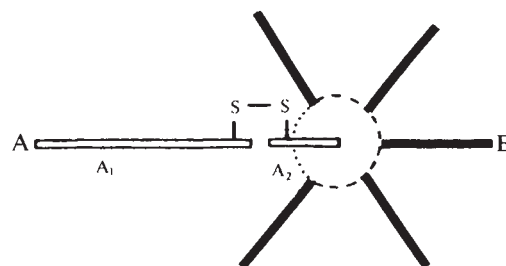


Fig. 1 Model of cholera toxin subunit structure. The toxin has one 28,000-MW A subunit and five 11,600-MW B subunits^{11-16,29,62,102-107}. The B subunits, which contain 103 residues 9 and 86 (refs 105, 106), are aggregated in a ring by tight, noncovalent bonds. The A subunit is linked to and partially inserted in the B ring through weaker noncovalent interactions⁴⁵. A subunit, although synthesized as a single polypeptide chain (one gene)¹⁰⁸, is usually 'nicked' between its two cysteine residues by bacterial protease(s) and thus splits into fragments A₁ (MW ~21,000) and A₂ (MW ~7,000) when treated with thiol-reducing agents^{11-15,62}. Reduction of whole toxin releases the A₁ fragment from the A₂-SB complex, indicating that A is attached to the B ring by its A₂ portion^{102,109}. Electron micrographs show the outer diameter of the B ring to be 90–100 Å, and the size of the separated A and B subunits as ~35 × 55 Å and ~24 × 30 Å (refs 110, 45).

Subsequent studies in several laboratories have provided further evidence that ganglioside G_{M1} is the natural biological receptor for cholera toxin.

(1) Studies of various cell types, including small intestinal mucosal cells of different species, have demonstrated a direct relationship between the cell content of G_{M1} and the number of toxin molecules that the cells can bind^{22,27,28}. In each cell type studied there have been about five molecules of G_{M1} per toxin binding site²⁷, which supports data from *in vitro* fixation studies²⁹⁻³² that each toxin B subunit binds to one G_{M1} molecule.

(2) Exogenous G_{M1} ganglioside can be incorporated into the cell membrane, there to act as a functional receptor. This was first shown by Cuatrecasas³³, who observed an increased cholera toxin-binding capacity and lipolytic responsiveness of fat cells which had been soaked in G_{M1} . Using ³H-labelled G_{M1} , J.H. *et al.*²² demonstrated the incorporation of G_{M1} into epithelial membrane of small intestine from humans and other species and showed that the increase in G_{M1} was associated with a corresponding increase in the capacity of the intestine to bind cholera toxin. *In vivo* tests in rabbits showed parallel increases of G_{M1} and susceptibility of the gut to the diarrhoeogenic action of the toxin²². Incorporation of G_{M1} into transformed cells deficient in this ganglioside has restored cell responsiveness to cholera toxin^{34,35}.

(3) Pretreatment of cell membranes with cholera toxin has been found to block specifically the membrane G_{M1} from reacting with galactose oxidase³⁶.

(4) Incubation of certain tissues with *V. cholerae* sialidase increases the number of toxin-binding sites in proportion to the additional G_{M1} produced by the enzyme from more complex gangliosides in the membrane: cellular sensitivity to the toxin is simultaneously enhanced^{27,113-115}. However, *V. cholerae* sialidase has failed to create new receptors for cholera toxin in intestinal epithelium²² even though after extraction, the intestinal gangliosides are normally hydrolysed by this enzyme²². Thus intestinal epithelium seems to possess a means of preserving its structural integrity from enzymatic attack.

(5) Chemical modifications of cholera toxin by various reagents have consistently and proportionally affected binding to cells and to plastic-adsorbed G_{M1} ganglioside³⁷.

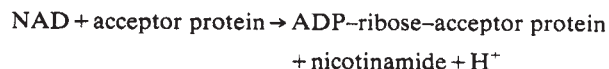
Membrane penetration and activation of adenylate cyclase

As already indicated, the lag between cell binding of cholera toxin and activation of adenylate cyclase mainly reflects the time taken for the toxin A subunit to penetrate the cell membrane. In comparison, the generation of A_1 and the subsequent effect on adenylate cyclase are rapid events that occur in <1 min (ref. 17). The cell penetration process is poorly understood; the time it takes is markedly influenced by the incubation temperature and the composition of the cell membrane (J.H. *et al.*, unpublished results). Lateral diffusion and capping of cholera toxin in the membrane have been demonstrated^{15,38-40} and have been proposed to be critical for action of the toxin^{38,41}. Lateral diffusion seems to have a much greater influence on the lag time in G_{M1} -poor cells than in those with more receptors. Thus, cell incorporation of exogenous G_{M1} significantly shortened the lag period in C6 cells which bind 7,000 toxin molecules per cell⁴², while no such effect was seen for mouse thymocytes (J.H. *et al.*, unpublished results) which bind 150,000 toxin molecules per cell and exhibit a much shorter lag period³⁷.

Moss *et al.*⁴³ and Tosteson and Tosteson⁴⁴ have shown that both cholera toxin and its B region can create pores in synthetic lipid bilayers containing G_{M1} ; however the observed capping of cholera toxin in viable cells suggests that the initial G_{M1} -toxin complex makes contact with an integral membrane protein. As discussed in detail elsewhere⁴⁵, the association of toxin with the membrane G_{M1} receptors might induce a conformational change in the B subunits, resulting in exposure of otherwise hidden hydrophobic B-subunit regions which then fuse with hydro-

phobic protein or lipid components of the plasma membrane. This could result in entry of the A subunit into the cell by one of at least three possible mechanisms. First, it is possible that the B subunits form a hydrophilic channel through the membrane, allowing A to pass; second, a channel for A subunit, may be built by the B subunits and an integral membrane protein; the third possibility is that a hydrophobic interaction between the A subunit and an integral membrane component may transpose A to the inner face of the membrane where intracellular glutathione would reduce the A_1 - A_2 disulphide linkage, thereby releasing A_1 . Membrane-bound cholera toxin is partly endocytosed^{27,46}. Recent studies have suggested that for diphtheria toxin (the action of which resembles that of cholera toxin although the metabolic and clinical effects are very different⁴⁷), adsorptive endocytosis is required for membrane penetration and cytotoxicity at physiological pH; lysosomotropic agents protect the cells, probably by neutralizing the acid pH in lysosomes, thus preventing the acid-dependent conformational change of endocytosed toxin in secondary lysosomes that is necessary for the translocation of the diphtheria toxin A fragment^{48,49}. Lysosomotropic agents, such as ammonium chloride and chloroquine, should also be tested for their effect on the activity of cholera toxin.

The intracellular biochemical events that lead to activation of adenylate cyclase are known. Gill⁵⁰ showed that, in broken cell preparations, activation depends on NAD, undefined cellular cytosol factors and ATP, in addition to the A_1 fragment and cell membrane. Moss *et al.*⁵¹ showed that cholera toxin, like diphtheria toxin, has ADP-ribosyltransferase activity, that is, it catalyses the reaction:



The protein that is ADP-ribosylated by cholera toxin has recently been identified as the guanyl nucleotide-binding component of the membrane-bound adenylate cyclase^{52,53}. Cassel and Selinger⁵⁴ have shown that adenylate cyclase is active while GTP is bound to the GTP-binding component but reverts to an inactive state as GTP is hydrolysed to GDP by GTPase; cholera toxin blocks the GTPase action which stabilizes adenylate cyclase in an active conformation. Alternatively, cholera toxin may stimulate adenylate cyclase by enhancing an exchange reaction in which stimulatory GTP replaces inhibitory GDP at a rate higher than that of hydrolysis of GTP to GDP⁵⁵.

B subunit cholera vaccine

Can recent knowledge about cholera toxin be used to improve prevention or treatment of disease? Although the development of the highly successful water and electrolyte substitution therapy was essentially an empirical process, three rather more specific approaches have given promising results in animals and are now being used in clinical trials.

The first of these is the development of an oral cholera vaccine based on enterotoxin B subunit. In contrast to clinical cholera disease, which gives rise to long-lasting immunity⁵⁶, the whole-cell cholera vaccines give only partial immunity for <6 months⁵⁷. Animal studies have shown that antitoxic and antibacterial cholera immunity cooperate synergistically in the gut giving a multiplicative protective effect by interfering with separate pathogenic events—toxin binding and bacterial adhesion and colonization⁵⁸⁻⁶⁰. The whole-cell vaccines probably fail both because they lack any toxin-derived antigen and because the injection route may be relatively inefficient in stimulating local immunity in the gut mucosa⁶¹.

Purified cholera B subunit, which spontaneously reassociates to the pentamer ring⁶⁰, is a logical 'toxoid' immunogen against cholera, especially for oral immunization. It is immunologically unrelated to A subunit^{13,29,62} and is a much stronger immunogen⁶³. Furthermore, isolated antibodies to B subunit have considerably higher cholera toxin-neutralizing activity than antibodies to the A subunit^{63,64}. The separation of B

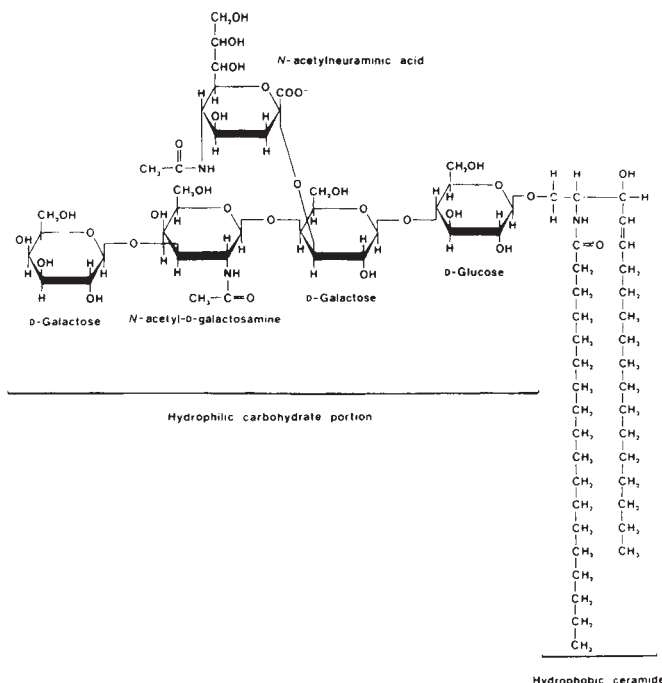


Fig. 2 Structure of the cholera toxin receptor, G_{M1} ganglioside. The oligosaccharide moiety of G_{M1} carries the binding determinants for cholera toxin, and studies have shown that both the terminal galactose and the sialic acid residues, positioned as in G_{M1} , are required for receptor activity^{25,111}; the removal or masking of any of these residues effectively inhibits the cholera toxin binding activity²⁵. Decarboxylation of the sialic acid residue destroys the receptor activity of G_{M1} (ref. 30 and L. Svennerholm *et al.*, in preparation), while oxidation with sodium periodate or exchange of the *N*-acetyl group by *N*-glycolyl does not affect activity (ref. 34 and L. Svennerholm *et al.*, in preparation). The lipid moiety of G_{M1} is also important. A sufficiently long hydrocarbon chain is needed for stable fixation of cholera toxin^{31,112}. Lyso- G_{M1} , in which the fatty acid has been replaced by an acetyl group, has been shown both to have intact toxin-binding activity³¹ and to be even better than G_{M1} in promoting subsequent toxin-membrane interactions necessary for activation of adenylate cyclase³⁴.

subunit from A excludes any risk of reversion to toxicity but does not lead to significant loss of protective antigen determinants⁶⁰. B subunit is particularly well suited to oral immunization because it retains the ability to bind to the intestinal epithelium, which has been shown to be important for stimulating mucosal immunity in animals, including local immunological memory⁶¹. Finally, as will be further discussed, B subunit given orally may also provide nonimmunological protection by blocking receptors before stimulating local immunoglobulin A (IgA) antitoxin formation. Purified B subunit is a strong protective antigen against experimental cholera in rabbits when given either alone or in combination with somatic antigens⁶⁰. It was recently tested⁶⁵ for its ability to stimulate mucosal immunity in humans; a single oral administration stimulated a marked local secretory IgA antibody response in 80% of the recipients. The response was comparable to that evoked in cholera patients by the natural disease⁶⁶. In further studies, two oral immunizations with a mixture of B subunit and whole-cell vaccine have produced secretory IgA antibody responses in intestine both to the toxin and the bacterial cell-wall lipopolysaccharide antigens in almost all people vaccinated⁶⁷. These results are promising but protective efficacy of an oral combined B subunit-whole cell vaccine in people living in areas in which cholera is endemic will require evaluation by field trials. The production of sufficient amounts of B subunit should not pose any problem, due to the recent development of affinity chromatography purification using a G_{M1} column⁶⁸.

Genetic methods may also supply *V. cholerae* strains which selectively lack the gene for the 'toxic' A subunit. By retaining the ability to colonize the intestine and produce immunogenic B subunit, such strains could be useful as live oral cholera vaccines. It must be borne in mind, however, that any living vaccine of this sort will require thorough testing for stability, and lack of side

effects before it can be used in humans. In that respect it is disappointing that the first such *V. cholerae* strain (Texas star)⁶⁹, produced diarrhoea (although usually very mild) in about 20% of recipient American volunteers in spite of producing no detectable holotoxin when grown *in vitro*⁷⁰.

Receptor-specific interference with toxin binding

The second rational approach to cholera prophylaxis is based on the identification of G_{M1} as the toxin receptor. In theory, the oral administration of large amounts of G_{M1} or a structural analogue could be used to prevent binding of the toxin to cell receptors. Alternatively, cell receptors could be blocked by the nontoxic B subunit.

Development of cholera was prevented by giving G_{M1} to rabbits³⁰, but for prophylactic medical use, enough G_{M1} would have to be given to ensure an excess in relation to the amount of toxin produced by the cholera vibrios in the gut. Most clinical isolates of *V. cholerae* produce considerably less than $1 \mu\text{g ml}^{-1}$ of cholera toxin when grown *in vitro*, and toxin concentrations in diarrhoeal fluid of patients have not been found to exceed $0.2 \mu\text{g ml}^{-1}$ (refs 71, 116). Assuming a daily production of $\sim 10 \text{ nmol}$ of cholera toxin in $\sim 5 \text{ l}$ of small intestinal juice, neutralization of the toxin should be accomplished by a similar amount of G_{M1} . To avoid the incorporation of oral G_{M1} into intestinal cells, which would increase their susceptibility to cholera²², the toxin could be adsorbed onto medical charcoal⁷¹, or covalently coupled to cellulose powder, for example (L. Svennerholm *et al.*, in preparation).

As G_{M1} cannot deactivate toxin that has already bound to the intestine, it would be expected to have a preventive rather than a curative effect. However, in a recent clinical trial the amount of G_{M1} -charcoal that completely bound the free, luminal *V. cholerae* enterotoxin also reduced purging in the early stage of disease⁷¹. The most likely explanation for this observation is that, because of the constant and rapid regeneration of gastrointestinal cells, there are always new cells available to bind cholera toxin or to be protected by an agent that binds the toxin. This is probably the first instance in which a specific receptor has been used to interfere with an infectious disease. However, it should be emphasized that the observed effect of G_{M1} -charcoal was too transient and incomplete to be practically useful; rather, G_{M1} given orally would possibly be useful for prophylaxis in high-risk groups such as family contacts of cholera patients.

As the B subunit of cholera toxin binds tightly to G_{M1} receptors but has no toxic activity, it might be possible to block the intestinal receptors by occupying them with purified B subunit protomer. It has been shown that pretreatment of the gut of rabbits with $0.5 \mu\text{g}$ of B subunit per cm completely protected the animals from experimental cholera after challenge with high doses of active cholera toxin^{22,72,73}. A human trial is conceivable now that it is technically possible to prepare purified B subunit in sufficient quantities⁶⁸; as little as $\sim 100 \mu\text{g}$ of B subunit could theoretically block all available receptors, based on the determination of amounts of G_{M1} in human intestinal epithelium²². B subunit is both safe and specific, and might prevent manifestation of disease while inducing local antitoxic immunity and allowing the infection to elicit an antibacterial immune response. In contrast to prophylactic G_{M1} , B subunit would also block toxin which is secreted in close proximity to the microvilli.

Antisecretory drugs

The third approach, intended for therapy rather than prophylaxis, aims at specifically reversing the toxic action after the cholera toxin has become bound to the intestinal cells. During the past decade oral hydration therapy by means of appropriate glucose-electrolyte solutions has greatly simplified the treatment of dehydrating diarrhoea, making effective therapy feasible in situations where intravenous treatment facilities are limited or unavailable⁷⁴⁻⁷⁶. This treatment takes advantage of the presence of an uptake mechanism for sodium in conjunction

with certain organic solutes, including glucose, which is not regulated by cyclic nucleotides and thus is unaffected in cholera and other forms of enterotoxic diarrhoea⁷⁷. In patients with mild or moderate dehydration the success rate with oral therapy is very high, usually more than 90%. However, in severely affected patients purging often occurs at such a high rate that balance cannot be maintained by oral fluid only, and intravenous replacement of fluid is essential for survival. Drugs that could reduce the rate of fluid loss in the severe dehydrating diarrhoeal diseases would obviate the need for much intravenous therapy and thus could play an important part in the management of severe diarrhoea. Such antisecretory drugs would need a high therapeutic index so that they could be used without extensive medical supervision, and they would have to be compatible with oral fluid therapy.

Based on present understanding of the biochemical events initiated by cholera toxin, possibilities for specific intervention include: (1) prevention of entry of A subunit (although not clinically useful as this event occurs before the onset of symptoms); (2) use of purines and their analogues to alter the NADase action of cholera toxin; (3) use of alternative acceptors, such as arginine or imidazoles, for the ADP-ribose generated by the toxin; (4) reversal of the ADP-ribose-adenylate cyclase association; and (5) removal of ADP-ribose by a specific ADP-ribosidase. While some of these approaches have reversed the action of cholera toxin *in vitro* in broken cell systems, kinetic and other considerations make the likelihood of success of inhibiting this stage of the secretory process *in vivo* extremely remote⁷⁸.

There seem to be considerably greater opportunities for inhibiting the action of cholera toxin subsequent to the ADP-ribosylation step by interfering with cyclic AMP formation or metabolism, or by modulating electrolyte-translocating mechanisms of the intestinal membrane. Several drugs have been found to inhibit secretion induced by cholera toxin in experimental animals, for example, chlorpromazine⁷⁹ and certain related compounds⁸⁰, nicotinic acid⁸¹, aspirin⁸², indomethacin⁸³, ethacrynic acid⁸⁴, propranolol⁸⁵, lidocaine⁸⁶ and berberine¹¹⁷ (for review see ref. 118). Chlorpromazine is the best studied of these drugs. In mice, treatment intramuscularly or enterally with 1–4 mg of chlorpromazine per kg body weight completely inhibited the intestinal secretion caused by cholera toxin, *Escherichia coli* LT, prostaglandin E₁ or dibutyryl cyclic AMP⁷⁹. In similar doses, chlorpromazine reversed fluid loss in piglets with enterotoxinogenic *E. coli* diarrhoea⁸⁷. In adult patients with severe cholera, treatment with chlorpromazine either perorally or intramuscularly, in a single dose of 1 or 4 mg per kg, rapidly and drastically reduced the purging volumes by ~65% (ref. 88). The patients became mildly sedated, were more comfortable, and had no nausea or vomiting. A follow-up clinical trial has shown that treatment of children with 1 mg per kg chlorpromazine increased the success rate of oral hydration therapy in patients with severe cholera⁸⁹.

The mechanism of antisecretory action of chlorpromazine is not fully known; however, it does seem to be confined to the intestinal epithelium^{79,90,91}. Chlorpromazine was selected for testing because of its ability to inhibit hormonal stimulation of cyclic AMP formation in various tissues^{92,93}; both the cholera toxin- and fluoride-stimulated adenylate cyclase activities of the intestinal mucosal membrane of chlorpromazine-treated mice are suppressed⁷⁹. However, the activity of the protein kinase of mucosal membranes is also reduced, suggesting multiple effects of chlorpromazine⁷⁹. In the brain, chlorpromazine and other phenothiazines are known to inactivate calmodulin and to interfere with transfer of calcium across the nerve-cell membrane⁹⁴; if these effects occur in the intestine, they could lead to inhibition of secretion^{91,119}. The action of chlorpromazine on calmodulin could possibly be linked to an inhibitory effect on adenylate cyclase.

Perspective

It has been estimated that each year about 1,000 million episodes of acute diarrhoea occur in children under 5 years of

age in Asia, Africa and Latin America, resulting in 5 million deaths. In many developing countries one-third to one-half of infant mortality can be attributed to diarrhoeal diseases. Even in the older age groups diarrhoeal episodes are frequent causes of severe illness and death^{95,96}.

Cholera is not the most prevalent of the diarrhoeal diseases, but it causes the most severe fluid loss and thus is responsible for a large proportion of life-threatening illness and death during the cholera season in endemic areas. Furthermore, research during the past decade has also made it clear that cholera is a prototype of diarrhoeal diseases caused by other enterotoxin-producing bacteria, especially *E. coli*⁹⁷. Taken together, the enterotoxic enteropathies probably account for at least one-half of dehydrating diarrhoeal illnesses. Knowledge of the structure and function of cholera toxin has been the main stimulus and provided guidelines for recent research in elucidating the aetiology and pathogenesis of these related diarrhoeas. Diagnostic assays for *E. coli* LT based on toxin-induced cyclic AMP manifestations in animals or cell lines^{98,99}, or more recently, on immunological¹⁰⁰ or receptor-immunological methods¹⁰¹ are based on similar methods for cholera toxin.

Although clean water supplies and safe disposal of human sewage would effect a dramatic reduction in cholera and other diarrhoeal disease, these two goals cannot be achieved in many areas for decades. If effective vaccines could be developed, they would no doubt be important components of national control programmes against diarrhoeal disease. The same holds true for drugs that could prevent or reduce fluid loss in enteric infections and thus alleviate the need for intravenous treatment.

Here I have described some theoretically possible approaches to counteract cholera toxin by immunological or pharmacological means. While each of these methods have been effective against cholera in animal studies, and, to the limited extent to which they have been tested, have also given promising results in man, much more research is needed to define their place, if any, in future prophylaxis and treatment of cholera. They should be regarded merely as starting points for further development. For example, in relation to vaccine development there is a need for better understanding of factors regulating the magnitude and duration of intestinal antitoxic as well as antibacterial immunity. With regard to receptor-prophylactic agents, basic research should be directed to identify specific binding or blocking agents for cholera vibrios which might cooperate synergistically with those for toxin. In the case of antisecretory agents, the aim should be to find drugs which effectively depress secretion but lack the sedative action of chlorpromazine. However, it is gratifying that cholera toxin research has now reached the stage where rational counteractive methods can be both formulated on the basic molecular knowledge, and, most importantly, be evaluated in controlled clinical and field trials.

I thank the many colleagues who contributed much to the cited work from this laboratory. I particularly thank Professor Lars Svennerholm and his group for close collaboration on all aspects of G_{M1} ganglioside as receptor for cholera toxin; Drs Ann-Mari Svennerholm and Ivar Lonnroth for aspects relating to immunity and vaccine development, and anti-diarrhoeal action of phenothiazines, respectively; Dr William B. Greenough and his staff at the International Centre for Diarrhoeal Disease Research, Dacca for the clinical trials; Institut Merieux, Lyon and SBL, Stockholm, for technical collaboration and for preparing G_{M1}-charcoal and B subunit for the clinical tests; and Dr Charles Wadsworth for his criticisms of the manuscript and helpful comments.

1. Ouchterlony, O. & Holmgren, J. (eds) *Cholera and Related Diarrhoeas: 43rd Nobel Symp.*, Stockholm 1978 (Karger, Basel, 1980).
2. De, S. N. *Nature* **183**, 1533–1534 (1959).
3. Dutta, N. K., Panse, M. V. & Kulkarni, D. R. *J. Bact.* **78**, 594–595 (1959).
4. Finkelstein, R. A. & LoSpalluto, J. J. *J. exp. Med.* **130**, 185–202 (1969).
5. Richardson, S. H., Evans, D. G. & Feeley, J. C. *Infect. Immunity* **1**, 546–553 (1970).
6. Field, M., Fromm, D., Wallace, C. K. & Greenough, W. B. *J. clin. Invest.* **48**, 24a (1969).
7. Schafer, D. E., Lust, W. D., Sircar, B. & Goldberg, N. D. *Proc. natn. Acad. Sci. U.S.A.* **67**, 851–856 (1970).
8. Sharp, G. W. G. & Hynie, S. *Nature* **229**, 266–269 (1971).
9. Field, M., Fromm, D., Al-Awqati, Q. & Greenough, W. B. *J. clin. Invest.* **51**, 795–804 (1972).

10. Greenough, W. B., Pierce, N. F. & Vaughan, M. *J. infect. Dis.* **121**, S111–S114 (1970).
11. Lönnroth, I. & Holmgren, J. *J. gen. Microbiol.* **76**, 417–427 (1973).
12. Cuatrecasas, P., Parikh, I. & Hollenberg, M. D. *Biochemistry* **12**, 4253–4264 (1973).
13. Holmgren, J. *Proc. 9th Joint US–Japan Cholera Conf.*, Grand Canyon 1973, 196–213 (Dept of State Publ. 8762, Washington, 1973).
14. Heyning, S. van. *Science* **183**, 656–657 (1974).
15. Holmgren, J., Lindholm, L. & Lönnroth, I. *J. exp. Med.* **139**, 801–819 (1974).
16. Finkelstein, R. A., Boesman, M., Neoh, S. H., LaRue, N. K. & Delaney, R. *J. Immun.* **113**, 175–180 (1974).
17. Gill, D. M. & King, C. A. *J. biol. Chem.* **250**, 424–432 (1975).
18. Heyning, S. van & King, C. A. *Biochem. J.* **146**, 269–271 (1975).
19. Flores, J. & Sharp, S. W. G. *J. clin. Invest.* **56**, 1345–1349 (1975).
20. Mekelanos, J. J., Collier, R. J. & Romig, W. R. *J. biol. Chem.* **254**, 5855–5861 (1979).
21. Cuatrecasas, P. *Biochemistry* **12**, 3547–3558 (1973).
22. Holmgren, J., Lönnroth, I., Månsson, J. E. & Svennerholm, L. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2520–2524 (1975).
23. Fishman, P. H. & Attkan, E. E. *J. biol. Chem.* **254**, 4342–4344 (1979).
24. Heyning, W. E. van, Carpenter, C. C. J., Pierce, N. F. & Greenough, W. B. *J. infect. Dis.* **124**, 415–418 (1971).
25. Holmgren, J., Lönnroth, I. & Svennerholm, L. *Infect. Immunity* **8**, 208–214 (1973).
26. King, C. A. & Heyning, W. E. van. *J. infect. Dis.* **127**, 639–647 (1973).
27. Hansson, H.-A., Holmgren, J. & Svennerholm, L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3782–3786 (1977).
28. Fishman, P. H. & Attkan, E. E. *J. biol. Chem.* **254**, 4342–4344 (1979).
29. Holmgren, J. & Lönnroth, I. *J. gen. Microbiol.* **86**, 49–65 (1975).
30. Sattler, J., Schwarzmann, G., Staerk, J., Ziegler, W. & Wiegand, H. *Hoppe-Seyler's Z. physiol. Chem.* **358**, 159–163 (1977).
31. Holmgren, J., Månsson, J.-E. & Svennerholm, L. *Med. Biol.* **52**, 229–233 (1974).
32. Fishman, P. H., Moss, J. & Osborne, J. C. *Biochemistry* **17**, 711–716 (1978).
33. Cuatrecasas, P. *Biochemistry* **12**, 3558–3566 (1973).
34. Fishman, P. H., Pacuska, T., Holm, B. & Moss, J. *J. biol. Chem.* **255**, 7657–7664 (1980).
35. Moss, J., Fishman, P. H., Manganiello, V. O., Vaughan, M. & Brady, R. O. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1034–1037 (1976).
36. Mullin, B. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **73**, 1679–1683 (1976).
37. Holmgren, J. & Lönnroth, I. *J. infect. Dis.* **133**, 564–574 (1977).
38. Craig, S. W. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3844–3848 (1975).
39. Revesz, T. & Greaves, M. *Nature* **257**, 103–106 (1975).
40. Sedlacek, H. H., Staerk, J., Seiler, F. R., Ziegler, W. & Wiegand, H. *FEBS Lett.* **61**, 272–276 (1976).
41. Bennet, V., O'Keefe, E. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 33–37 (1975).
42. Fishman, P. H. in *Secretory Diarrhea* (eds Field, M., Fordtran, J. S. & Schultz, S. G.) 85–106 (American Physiological Society, Maryland, 1980).
43. Moss, J., Richards, R. L., Alving, C. R. & Fishman, P. H. *J. biol. Chem.* **252**, 797–798 (1977).
44. Tosteson, M. T. & Tosteson, D. C. *Nature* **275**, 142–144 (1978).
45. Holmgren, J. & Lönnroth, I. in *Cholera and Related Diarrhoeas: 43rd Nobel Symp.*, Stockholm 1978, 88–103 (Karger, Basel, 1980).
46. Manuelidis, L. & Manuelidis, E. E. *J. Neurocytol.* **5**, 575–589 (1976).
47. Pappenheimer, A. M. Jr. *Rev. Biochem.* **46**, 69–94 (1977).
48. Sandvig, K. & Olsnes, S. *J. Cell Biol.* **87**, 828–832 (1980).
49. Draper, R. K. & Simon, M. I. *J. Cell Biol.* **87**, 849–854 (1980).
50. Gill, D. M. *J. infect. Dis.* **133**, S55–S63 (1976).
51. Moss, J., Manganiello, V. C. & Vaughan, M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4424–4427 (1976).
52. Cassel, D. & Pfeuffer, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2669–2673 (1978).
53. Gill, D. M. & Mehren, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3050–3054 (1978).
54. Cassel, D. & Selinger, Z. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3307–3311 (1977).
55. Lad, P. M., Nielsen, T. B., Preston, N. S. & Rodbell, M. *J. biol. Chem.* **255**, 988–995 (1980).
56. Cash, R. A. *et al. J. infect. Dis.* **130**, 325–333 (1974).
57. Feeley, J. C. & Gangarosa, E. J. in *Cholera and Related Diarrhoeas: 43rd Nobel Symp.*, Stockholm 1978, 204–210 (Karger, Basel, 1980).
58. Svennerholm, A.-M. & Holmgren, J. *Infect. Immunity* **13**, 735–740 (1976).
59. Peterson, J. W. *Infect. Immunity* **26**, 528–533 (1979).
60. Holmgren, J. *et al. Nature* **269**, 602–604 (1977).
61. Pierce, N. F. *J. exp. Med.* **148**, 195–206 (1978).
62. Ohtomo, N., Muraoka, T. & Tashiro, A. *Proc. 10th Joint US–Japan Cholera Conf.*, Kyoto 1974, 115–118 (Tokyo, 1974).
63. Svennerholm, A.-M. in *Cholera and Related Diarrhoeas: 43rd Nobel Symp.*, Stockholm 1978, 171–184 (Karger, Basel, 1980).
64. Peterson, J. W., Hejtmancik, K. E., Markel, D. E., Craig, J. P. & Kurosky, A. *Infect. Immunity* **24**, 774–779 (1979).
65. Svennerholm, A.-M., Sack, D. A., Bardhan, P. K., Jertborn, M. & Holmgren, J. in *Proc. 16th US–Japan Cholera Conf.*, Gifu, 1980 (Osaka University Press, in the press).
66. Sack, D. A., Islam, A., Holmgren, J. & Svennerholm, A.-M. in *Proc. 15th Joint US–Japan Cholera Conf.*, Bethesda, 1979, 423–439 (NIH Publ. 80-2003, 1980).
67. Svennerholm, A.-M., Jertborn, M., Gothefors, L., Sack, D. A. & Holmgren, J. (in preparation).
68. Tayot, J.-L., Holmgren, J., Svennerholm, L., Lindblad, M. & Tardy, J. *Eur. J. Biochem.* **113**, 249–258 (1981).
69. Honda, T. & Finkelstein, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2052–2056 (1979).
70. Levine, M. in *Nobel Conf. 3* (eds Holme, T., Holmgren, J., Merson, M. & Möllby, R.) (Elsevier, Amsterdam, in the press).
71. Stoll, B. J. *et al. Lancet* **ii**, 888–891 (1980).
72. Pierce, N. F. *J. exp. Med.* **137**, 1009–1023 (1973).
73. Holmgren, J. *Infect. Immunity* **8**, 851–859 (1973).
74. Hirschhorn, N. J. *et al. New Engl. J. med.* **279**, 176–184 (1968).
75. Nalin, D. R., Cash, R. A., Islam, R., Molla, M. & Phillips, R. A. *Lancet* **ii**, 370–372 (1968).
76. Pierce, N. F. *et al. Ann. intern. Med.* **70**, 1173–1183 (1969).
77. Schultz, S. G. & Curran, P. F. *Curr. Topics Membrane Transport* **5**, 225–281 (1974).
78. Greenough, W. B. III & Hendrix, T. R. in *Cholera and Related Diarrhoeas: 43rd Nobel Symp.*, Stockholm 1978, 234–244 (Karger, Basel, 1980).
79. Holmgren, J., Lange, S. & Lönnroth, I. *Gastroenterology* **75**, 1103–1108 (1978).
80. Lönnroth, I., Lange, S. & Holmgren, J. in *Phenothiazines and Structurally Related Drugs: Basic and Clinical Studies* (eds Usdin, E., Eckert, H. & Forrest, S. J. J.) 303–306 (Elsevier, New York, 1980).
81. Turjman, N., Gotterer, G. S. & Hendrix, T. R. *J. clin. Invest.* **61**, 1155–1160 (1978).
82. Finch, A. D. & Katz, R. L. *Nature* **238**, 273–274 (1972).
83. Jacoby, H. I. & Marshall, C. H. *Nature* **235**, 163–165 (1972).
84. Carpenter, C. C. J., Curlin, G. T. & Greenough, W. B. III *J. infect. Dis.* **120**, 332–338 (1969).
85. Taub, M., Bonorris, G., Chung, A., Coyne, M. J. & Schoenfeld, L. J. *Gastroenterology* **72**, 101–105 (1977).
86. Cassuto, J., Jodal, M., Tuttle, R. & Lundgren, O. *Experientia* **35**, 1467–1468 (1979).
87. Lönnroth, I., Andrén, B., Lange, S., Martinsson, K. & Holmgren, J. *Infect. Immunity* **24**, 900–905 (1979).
88. Rabbani, G. H., Greenough, W. B. III, Holmgren, J. & Lönnroth, I. *Lancet* **i**, 410–412 (1979).
89. Islam, M. R., Sack, D. A., Holmgren, J., Bardhan, P. K. & Rabbani, *Lancet* (submitted).
90. Lönnroth, I. & Munck, B. G. *Acta pharmac. tox.* **47**, 439–446 (1979).
91. Smith, P. L. & Field, M. *Gastroenterology* **78**, 1545–1553 (1980).
92. Kakiuchi, S. & Rall, T. W. *Molec. Pharmac.* **4**, 367–378 (1968).
93. Wolf, J. & Jones, A. B. *Proc. natn. Acad. Sci. U.S.A.* **65**, 454–459 (1970).
94. Cheung, W. Y., Lynch, T. J. & Wallace, R. W. *Adv. Cyclic Nucleotide Res.* **9**, 233–251 (1978).
95. Mata, L., Kronmal, R. A. & Villegas, H. in *Cholera and Related Diarrhoeas: 43rd Nobel Symp.*, Stockholm 1978, 1–14 (Karger, Basel, 1980).
96. Barua, D. in *Nobel Conf. 3* (eds Holme, T., Holmgren, J., Merson, M. & Möllby, R.) (Elsevier, Amsterdam, in the press).
97. Sack, R. B. A. *Rev. Microbiol.* **29**, 333–353 (1975).
98. Donta, S. T. & Smith, D. M. *Infect. Immunity* **9**, 500–505 (1974).
99. Guerrant, R. L., Brunton, L. L., Schnaitman, T. C., Rebhun, L. I. & Gilman, A. G. *Infect. Immunity* **10**, 320–327 (1974).
100. Yolken, R., Greenberg, H. B., Merson, M., Sack, R. B. & Kapikian, A. Z. *J. clin. Microbiol.* **6**, 439–444 (1977).
101. Svennerholm, A.-M. & Holmgren, J. *Curr. Microbiol.* **1**, 19–27 (1978).
102. Sattler, J. *et al. Eur. J. Biochem.* **57**, 309–316 (1975).
103. Gill, D. M. *Biochemistry* **15**, 1242–1248 (1976).
104. Lai, C.-Y., Mendez, E. & Chang, D. J. *J. infect. Dis.* **133**, S23–S30 (1976).
105. Kurosky, A., Markel, D. E. & Peterson, J. W. *J. biol. Chem.* **252**, 7257–7264 (1977).
106. Lai, C.-Y. *J. biol. Chem.* **252**, 7249–7256 (1977).
107. Sigler, P. B., Druyan, M. E., Kiefer, H. C. & Finkelstein, R. A. *Science* **197**, 1277–1279 (1977).
108. Gill, D. M. & Rappaport, R. S. *J. infect. Dis.* **139**, 674–680 (1979).
109. Tomasi, M., Battistini, A., Araco, A., Roda, L. G. & D'Agnolo, G. *Eur. J. Biochem.* **93**, 621–627 (1979).
110. Ohtomo, N., Muraoka, T., Tashiro, A., Zinnaka, Y. & Amako, K. *J. infect. Dis.* **133**, (Suppl.), S31–S40 (1976).
111. Fishman, P. J., Moss, J., Richards, R. L., Brady, R. O. & Alving, C. R. *Biochemistry* **18**, 2562–2567 (1979).
112. Staerk, J., Ronneberger, H. J., Wiegand, H. & Ziegler, W. *Eur. J. Biochem.* **48**, 103–110 (1974).
113. Haksar, A., Maudsley, D. V. & Péron, F. G. *Nature* **251**, 514–515 (1974).
114. Révész, T., Greaves, M. F., Capellaro, D. & Murray, R. K. *Br. J. Haemat.* **34**, 623–630 (1976).
115. Holmgren, J. in *Bacterial Toxins and Cell Membranes* (eds Jeljaszewicz, I. & Wadström, T.) 333–366 (Academic, London, 1978).
116. Craig, J. P. *Nature* **207**, 614–616 (1965).
117. Sack, R. B. in *Nobel Conf. 3* (eds Holme, T., Holmgren, J., Merson, M. & Möllby, R.) (Elsevier, Amsterdam, in the press).
118. Powell, D. W. & Field, M. in *Secretory Diarrhoea* (eds Field, M., Fordtran, J. S. & Schultz, S. G.) 187–209 (American Physiological Society, Maryland, 1980).
119. Jlundain, A. & Naftalin, R. J. *Nature* **279**, 446–448 (1979).

ARTICLES

High noble metal concentrations in a late Pliocene sediment

Frank T. Kyte[†], Zhiming Zhou^{*} & John T. Wasson^{††}

^{*}Institute of Geophysics and Planetary Physics, [†]Department of Earth and Space Sciences, ^{††}Department of Chemistry, University of California, Los Angeles, California 90024, USA

A 2.3-Myr-old layer in a sediment from the Antarctic Ocean contains Ir and Au at levels comparable with those at the Cretaceous–Tertiary boundary. A sizable fraction of the noble metals is contained in vesicular, millimetre-sized poly-mineralic grains that closely resemble ablation debris from chondritic meteorites, and there is little doubt that the noble metals resulted from the accretion of a large extraterrestrial object. No massive extinctions or other evidence of environmental stress seem to be associated with this accretionary event.

THE accretion of kilometre-sized asteroids to the Earth must have a significant impact on the terrestrial ecosystem. Such encounters have been a stochastic source of stress to the biosphere, influencing evolution and extinction rates. Recent discoveries of Ir (ref. 1) and other extraterrestrial siderophiles¹⁻⁴ in sediments at the 65-Myr-old Cretaceous-Tertiary boundary have supported the hypothesis that encounters of sufficient magnitude can lead to the simultaneous extinction of many species. These discoveries have also illustrated that at least some types of encounters deposit a measurable geochemical signal in the sedimentary record.

If the Earth accreted a massive body at the end of the Cretaceous 65-Myr ago, then other bodies of similar size should have left their signatures elsewhere in the geological record. From relationships given by Grieve and Dence⁵ and an assumed mean impact velocity of 20 km s^{-1} we calculate the relationship $N = 6.7 \times 10^{-7} r^{-2}$ where N is the cumulative number of bodies impacting the Earth each year and r (in km) is the radius of the body. This relationship indicates that ~ 43 bodies should have impacted the Earth since the Cretaceous. Wetherill⁶ gives a similar estimate. This relationship is based on crater counts, and if preimpact breakup occurs in an appreciable fraction of the cases, the impact rate could be higher⁷.

A logical step in establishing the link between major changes in the terrestrial environment and the accretion of kilometre-sized bodies is to find and characterize other sediments having enhanced concentrations of noble metals. In this article we report our first success, the discovery a large enhancement of noble metals near the Pliocene-Pleistocene boundary in an Antarctic Ocean deep-sea core. Core USNS Eltanin 13-3 was selected because Crockett and Kuo⁸ found Ir concentrations in one 210-cm long section that were three times higher than those in other sections from the same core. It was also intriguing that the base of this 210-cm unit was associated with radiolarian extinctions, geomagnetic reversals, and climatic change⁹.

Experimental results

Eltanin core 13-3 is a muddy diatomaceous ooze collected in 1964 at a depth of 5,090 m at $57^\circ 00' \text{ S}$ and $89^\circ 29' \text{ W}$. It has been heavily sampled and is dried out in some areas. In the region of interest (780–800 cm) it contains $\sim 65\%$ diatoms and radiolaria and 30% clay with $\sim 5\%$ quartz, heavy minerals and trace Mn micronodules. It consists of a pale brown ooze with strings or sedimentary clasts of pale orange diatomaceous ooze. The sedimentation rate as determined by palaeomagnetic stratigraphy is approximately constant at $3.3 \mu\text{m yr}^{-1}$ (refs 8, 9). A total of 71 samples, each 3–4 cm in length were taken for a continuous section from 645 to 875 cm, extending from 10 cm above to 10 cm below the interval in which Crockett and Kuo⁸ found the noble-metal enhancement. Two additional control samples were taken from 551–553 and 946–948 cm. Samples of 150–200 mg were irradiated for 60 h at $10^{14} \text{ neutrons cm}^{-2} \text{ s}^{-1}$ and analysed by instrumental neutron activation analysis (INAA); results are listed in Table 1. Iridium was the only noble metal that could be determined; the control samples and those in the ranges 645–771 and 797–875 mm were below our detection limit of 0.5 ng per g. On discovery of high Ir (1–9 ng per g) levels in samples from the 771–797 cm level, another 14 samples (600–750 mg) were irradiated for 3 h at a flux at $2 \times 10^{12} \text{ neutrons cm}^{-2} \text{ s}^{-1}$ in the UCLA reactor and Au and Ir were determined by radiochemical neutron activation analysis (RNAA) using the procedure described previously⁴.

Our RNAA analyses on a second set of samples were carried out to confirm the Ir anomaly and to measure Au concentrations in this horizon (Table 1). We also attempted to analyse for Os, Pt, Pd and Re in two samples, but because of the low neutron flux and low noble metal concentrations we obtained only upper limits with CI-chondrite-normalized element/Ir ratios > 2 . These are not diagnostic for distinguishing between chondrite groups, but we plan to determine these elements in another experiment. Poor chemical yields caused significant errors on a few Ir data, but in general there was a good correlation between

Au and Ir. Our background Ir values agree well with the value of 0.20 ng per g determined by Crockett and Kuo⁸; our background Au value of $\sim 0.32 \text{ ng per g}$ is about half their value. We cannot readily explain this discrepancy. Our RNAA and INAA Ir values often differed by as much as a factor of 2. As discussed below, these discrepancies probably reflect sample inhomogeneity rather than analytical problems. The band proved wider in the RNAA section than anticipated on the basis of the initial INAA survey. Later, in longer counts an Ir concentration of 2.3 ng per g was found in another sample (35) which may extend the enriched layer another 7 cm above the RNAA data set.

Because the sample sizes were about four times larger we have greater confidence in our RNAA data. Figure 1 shows the RNAA Ir and Au concentrations in the peak and in several background samples. Using a background of 0.20 and 0.32 ng per g for Ir and Au, respectively, and a bulk dry density of the sediment of 2 g cm^{-3} , the net Ir and Au fluences integrated across the horizon are 100 and 63 ng cm^{-2} , respectively. The resulting Au/Ir mass ratio of 0.6 is higher than that in all groups of chondrites except the EH (formerly designated E4 and E5) enstatite chondrites¹⁰.

By chance, two large (1.5–2.0 mm) grains of Ir-rich material were found in INAA sample 42. Together with a minor amount of adhering clay they weighed 4 mg. Their Ir concentration was 160 ng per g, accounting for $\sim 40\%$ of the total Ir in this sample. In polished section these grains were observed to be highly vesicular ($\sim 50\%$) with a finely microcrystalline texture. This consisted of equant, $\sim 10 \mu\text{m}$ euhedral silicates in a silicate matrix containing a sparse submicrometre opaque phase. Qualitative analysis by electron microprobe indicated that the silicates were zoned Ni-bearing olivine in a Ca, Al, Si-rich matrix.

It seems that these materials correspond to the olivine-magnetite-glass assemblage typical of cosmic spherules and debris produced by the laboratory ablation of chondrites¹¹.

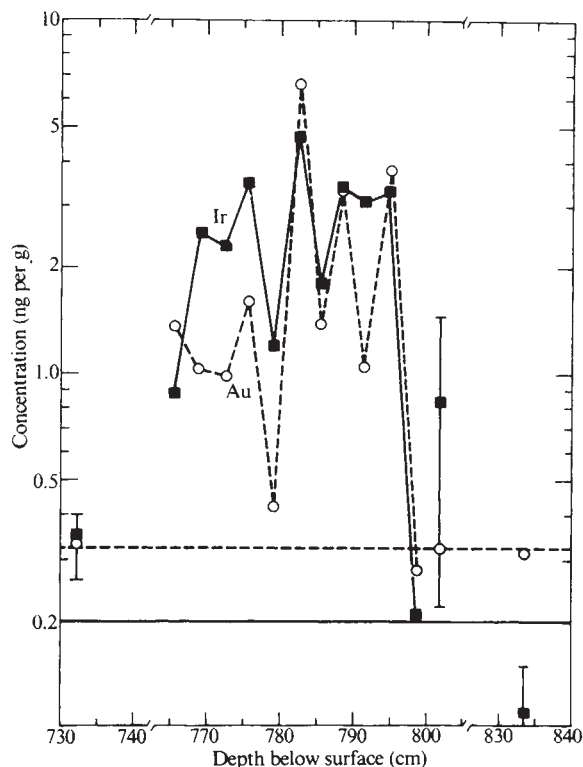


Fig. 1 Concentrations of Ir and Au in Eltanin core 13-3. Background levels (shown by horizontal lines) are 0.20 ng per g Ir and 0.32 ng per g Au; maximum concentrations in the peak in the depth range 765–800 cm are ~ 25 times the background levels. The integrated flux of Ir is 100 ng cm^{-2} , of Au 63 ng cm^{-2} . The observed Au/Ir ratio is similar to values in EH enstatite chondrites. The 70% uncertainty limits are shown for three of the lowest Ir values are examples of the precision at these levels. More exact information regarding the precision is given in Table 1.

Table 1 Concentrations of Ir and Au measured at various depths of Eltanin core 13-3

Sample	Depth (cm)	INAA	RNAA	
		Ir (ng per g)	Ir (ng per g)	Au (ng per g)
1-25	645-731	<0.5	—	—
27	731-734	≤0.5	0.33 ⁺	0.35
28-34	734-757	≤0.5	—	—
35	757-760	2.3	—	—
36	760-764	≤0.5	—	—
37	764-767	≤0.5	0.88	1.36
38	767-771	≤0.6	2.5	1.03
39	771-774	1.2	2.3	0.98
40	774-777	2.0	3.5*	1.6
41	777-781	6.7	1.2	0.42
42	781-784	10	4.7	6.6
43	784-787	1.4	1.8	1.39
44	787-790	1.3	3.4*	3.3
45	790-793	1.3	3.1	1.05
46	793-797	1.4	3.3	4.8
47	797-800	≤0.5	0.21*	0.28
48	800-804	≤0.5	0.83‡	0.32
49-57	804-832	≤0.5	—	—
58	832-835	≤0.5	0.11 ⁺	0.31
59-71	835-875	≤0.5	—	—

Relative uncertainty limits are $\pm \leq 10\%$ except as indicated.

* 70% confidence limits $\pm 10-19\%$.

† 70% confidence limits $\pm 20-39\%$.

‡ 70% confidence limits $\pm 40-79\%$.

Their texture, equant olivine grains set in a vesicular matrix, is found in cosmic spherules separated from sea sediments, but is less common than textures such as barred olivine or platy brickwork¹¹. A $>147\text{-}\mu\text{m}$ sieve fraction of another 1g of sample 42 contained a 1-mm particle with similar petrographic properties and four others in the size range 0.2–0.5 mm. The sample heterogeneity mentioned above is probably the result of statistical fluctuations in the number of large grains of chondritic debris through the section.

Only one particle was roughly spherical in outline, the rest having irregular shapes. All particles had severely etched rims, with only their cores retaining an apparently unaltered texture. The obvious dissolution of particle rims may have resulted in a post-depositional redistribution of a fraction of the noble metals in the horizon. Our observed concentrations of $\geq 3\text{ mg cm}^{-3}$ chondritic debris contrasts with typical influxes of ablation debris of 'cosmic spherules' $>0.2\text{ mm}$ of 30 ng cm^{-3} based on a worldwide accretion rate of 90 tons yr^{-1} (ref. 12) and the reported sedimentation rate.

Discussion

There seems little doubt that the noble metals in our Antarctic Ocean core are of extraterrestrial origin, as indicated by their roughly chondritic Au/Ir ratios, and by the remarkably high concentration of chondritic debris. This debris reaches concentrations about five orders of magnitude higher than expected in sediment having the observed sedimentation rate. The observed particles have petrographic textures of a rare type regularly observed in typical collections of cosmic spherules.

We interpret this horizon as resulting from a single, accretionary event. Presently available information does not allow us to distinguish between oceanic impact and atmospheric ablation as the source of these materials, but similarities of the large 'spherules' to materials produced in artificial ablation experiments¹³ suggests that they were produced by atmospheric ablation of asteroidal or cometary material. Total breakup above the solid or liquid surface of the Earth apparently requires that the material be highly friable as is inferred to be the case for some cometary materials. Microtektites have not been found in sieved samples of our horizon and thus are much less abundant than 'ablation' debris. This probably indicates that at most a minor fraction of the noble metal enhancement can be associated with crater ejecta.

The integrated flux of $80-100\text{ ng Ir cm}^{-2}$ is $\sim 0.15\text{ g cm}^{-2}$ $\text{H}_2\text{O-free CI material}$ or an initial horizon $\sim 0.5\text{ mm}$ thick. We suspect that bioturbation has spread out an initially thin, concentrated horizon, although this is difficult to confirm observationally in such an old, heavily sampled core. The 30–40 cm thickness of the enriched layer is similar to thicknesses of microtektite horizons¹⁴ and to bioturbation mixing lengths inferred in core E13-3 in palaeontological studies¹⁵. Rapidly deposited horizons of this type might, in fact, provide interesting markers for estimating effects of bioturbation and of trace-element transport during diagenesis.

The $80-100\text{ ng Ir cm}^{-2}$ in this horizon is similar to the maximum reported for the Cretaceous–Tertiary boundary at various European sites^{1,2,4}. Thus, if this horizon is worldwide, the mass of accreted chondritic matter is comparable with that of the Cretaceous–Tertiary materials. However, the high abundance of millimetre-sized ablation (or impact) debris suggests that the site of this core is near the location where the meteoroid entered the atmosphere and/or impacted with the surface of the ocean, thus the fluence at distant locations may be significantly lower.

The late Pliocene noble-metal-enrichment is probably not too localized, otherwise it would probably not have been discovered in the few cores analysed by Crockett and Kuo⁸. A major event showering debris over a wide area is a much more likely mechanism for producing discoverable concentrations of this magnitude. Considering that the proposed 0.5 mm-thick horizon was deposited after settling through 5 km of water, and the differential settling times of the particles ($\sim 8\text{ h}$ for 2 mm compared with $\sim 11\text{ h}$ for 1 mm diameters) in currents which today are $\sim 0.7\text{ km h}^{-1}$ (ref. 16) would spread this debris over a very large area even if it originated from a point source. We propose a minimum projectile diameter of 20 m, enough material to produce similar enrichments over only 10 km^2 . We suspect that it must have been significantly larger.

Crockett and Kuo⁸ found no Ir enhancement at the same time range in core Eltanin 17-10 that was collected $\sim 2,800\text{ km}$ away. Their sample containing the corresponding time interval, however, covers a depth range 800–1,400 cm, nearly three times longer than that in 13-3. For this reason their negative result may not be significant. Additionally, palaeomagnetic studies indicate that the horizon would be near the base of the sampled interval and might even overlap into the next interval (1,400–1,585 cm) which has slightly higher Ir (0.42 ng per g).

In contrast to noble metal-enriched sediments at the Cretaceous–Tertiary boundary, there is no major palaeoenvironmental crisis associated with this noble metal horizon. Hayes and Opdyke⁹ report that radiolarian species vary with depth in core 13-3, but no dramatic change is associated with the peak in Ir concentrations at the 770–800 cm depth. In contrast to the Cretaceous–Tertiary boundary, the nearby Pliocene–Pleistocene boundary does not represent a major biostratigraphic hiatus. Definition of the base of the Pleistocene is now generally accepted as the base of the Calabrian stage in the type section at La Castella, Italy, dated at $\sim 1.8\text{ Myr}$ (refs 17, 18). Earlier attempts to define this boundary based on the onset of glaciation (continental or polar) resulted in proposed boundaries ranging in age from 0.5 to 4 Myr (ref. 18); major cooling occurred as early as 4.0 Myr (ref. 17). Palaeomagnetic evidence⁹ indicates that our horizon occurs at $\sim 2.3\text{ Myr}$. If, in fact, this horizon proves to be worldwide it is an important marker bed for late Pliocene stratigraphy, and might warrant use as the Pliocene–Pleistocene boundary.

Clearly it is important to extend our studies to sediments of the same age from other localities. Although our best projection based on the available facts is that noble metal concentrations at distant locations will be less than those in core 13-3, we do expect to find that the horizon covers a significant fraction of the Earth. Additionally, the discovery of a second significant horizon during the most recent 65 Myr of Earth history suggests that we can find others of similar magnitude throughout the Cenozoic. If so, this will lead to major improvements both in our knowledge regarding the flux of large bodies to the Earth and in

our ability to correlate marine sediments from widely separated locations.

If this horizon does prove to be worldwide in extent, the absence of major extinctions associated with this late Pliocene event becomes an important fact bearing on models explaining the massive Cretaceous-Tertiary extinctions in terms of environmental effects associated with a major influx of extra-terrestrial matter.

Received 16 March; accepted 22 May 1981.

1. Alvarez, L. W., Alvarez, W., Asaro, F. & Michel, H. V. *Science* **208**, 1095-1108 (1980).
2. Smit, J. & Hertogen, J. *Nature* **285**, 198-200 (1980).
3. Ganapathy, R. *Science* **209**, 921-923 (1980).
4. Kyte, F. T., Zhou, Z. & Wasson, J. T. *Nature* **288**, 651-656 (1980).
5. Grieve, R. A. F. & Dence, M. R. *Icarus* **38**, 230-242 (1979).
6. Wetherill, G. W. *Scient. Am.* **240** (3), 54-65 (1979).
7. Melosh, H. J. & Passey, Q. *Proc. Conf. Multiring Basins* (Lunar Planetary Institute, Houston, 1981).
8. Crockett, J. H. & Kuo, H. Y. *Geochim. cosmochim. Acta* **43**, 831-842 (1979).
9. Hays, J. D. & Opdyke, N. D. *Science* **158**, 1001-1011 (1967).

We thank Dennis S. Cassidy for the Eltanin core, and H. Tappan, A. R. Loeblich and G. W. Wetherill for helpful comments. The core description was provided by F. A. Kaharoed-din, and technical assistance by J. Pai and A. Young. This research was supported in part by the Office of Naval Research. Curation of Eltanin cores at Florida State University is supported by NSF contract DPP75-19723.

10. Mason, B. *Handbook of Elemental Abundances in Meteorites* (Gordon and Breach, New York, 1971).
11. Blanchard, M. B., Brownlee, D. E., Bunch, T. E., Hodge, P. W. & Kyte, F. T. *Earth planet. Sci. Lett.* **46**, 178-190 (1980).
12. Murrell, M. T., Davis, P. A. Jr, Nishizumi, K. & Millard, H. T. Jr *Geochim. cosmochim. Acta* **44**, 2067-2074 (1980).
13. Kyte, F. T. M. S. thesis, San Jose State Univ. (1977).
14. Glass, B. P. & Zwart, M. J. *Bull. geol. Soc. Am.* **90**, 595-602 (1979).
15. Hayes, J. D. *Bull. geol. Soc. Am.* **82**, 2433-2447 (1971).
16. Goodell, H. G., Meylan, M. A. & Grant, B. *AGU Antarctic Res. Ser.* **15**, 27-92 (1971).
17. Berggren, W. A. *Micropaleontology* **24**, 337-370 (1978).
18. Berggren, W. A. & Van Couvering, J. A. in *Treatise on Invertebrate Paleontology, A* (eds Robinson, R. A. & Teichert, C.) 505-543 (Geological Society of America, New York, 1979).

Controls of RNA splicing and termination in the major late adenovirus transcription unit

Goran Akusjarvi & Hakan Persson

Department of Microbiology, University of Uppsala, The Biomedical Center, Box 581, S-751 23 Uppsala, Sweden

The major late adenovirus promoter is active early after infection, selectively producing messenger RNAs coding for polypeptides with molecular weights of 55,000, 52,000 and 14,000. This selective expression suggests that a differential splicing pattern occurs at the transition from early to late viral gene expression. Activation of the late promoter and splicing of the 55, 52K mRNAs does not require newly synthesized virus polypeptides.

TRANSCRIPTION from the adenovirus genome late in infection initiates predominantly at the promoter located at map coordinate 16.3 (ref. 1) and proceeds in a rightward direction for about 28,000 base pairs (bp)². This long viral transcript is processed into five families of messenger RNAs (mRNAs), each family having co-terminal 3' ends³⁻⁵ (L1-L5, Fig. 1A) and each mRNA containing a tripartite leader, encoded at coordinates 16.6, 19.6 and 26, spliced onto the body of the mRNA^{6,7}. The genes for two low-molecular-weight RNAs (VA RNA_I and VA RNA_{II}), transcribed by RNA polymerase III, have been located close to coordinate 30 on the viral genome^{8,9}. Three mRNA species, originating from the major late transcription unit, have been mapped by electron microscopy in the vicinity of the VA RNA genes¹⁰ (Fig. 1A).

The lytic cycle of adenovirus infection is divided into two functionally distinct phases, the early phase preceding and the late phase beginning with the onset of virus DNA replication. It has been assumed that the major late transcription unit is activated at the shift from early to late phase. However, a low level of rightward-reading transcripts proximal to the late promoter has been detected in the nuclei of early infected cells¹¹, and truncated cytoplasmic RNAs originating from the late promoter have also been found early in the infectious cycle by electron microscopy^{12,13}. Transcription studies using purified restriction fragments carrying the major late promoter suggest that virus-specific components are not necessary for initiation at this promoter site, at least *in vitro*^{14,15}. We now demonstrate that the major late transcription unit is active at early times of infection, producing selected transcripts. This indicates a novel, virus-regulated abundance control of mRNA production.

Viral RNAs from region L1

To study the expression of RNAs from the major late transcription unit, we first determined the structure of the r strand-specific RNAs encoded in region L1. Cytoplasmic RNA prepared late after adenovirus type 2 (Ad2) infection was hybridized to the entire viral DNA and treated with S₁ endonuclease. The S₁-resistant hybrids were separated on an alkaline agarose

gel and subsequently transferred to a nitrocellulose filter. Hybridization of the filter to a ³²P-labelled probe specific for region L1 (Fig. 2) shows that five major RNA species (Fig. 1B, lanes a-e) had been present in the cytoplasmic RNA preparation. Species b, c and e have previously been identified by electron microscopy¹⁰. The S₁ analysis revealed two novel species, a and d. The length of species a suggests that it represents a co-linear transcript of the DNA template from the third leader segment up to the poly(A) addition site at position 38.5 (ref. 3). Late viral cytoplasmic RNA was selected by hybridization to a BamHI-HindIII fragment specific for region L1. The selected RNA was then used to determine the structure of region L1 RNAs. The RNA was first hybridized to the HindIII-B fragment¹⁶ of the viral DNA (coordinates 17.0-31.5), ³²P-labelled by nick translation followed by S₁ endonuclease treatment¹⁷. This analysis revealed two major protected species, 520 and 440 nucleotides long (Fig. 2B, lane a). Further mapping of protected fragments showed that VA RNA_I, VA RNA_{II} and a 200-nucleotide read through product of the VA RNA_I gene (V₂₀₀)¹⁸ were present in the selected RNA. In the same way, the bands of 73 and 89 nucleotides were shown to correspond to the second and third segments of the common tripartite leader^{19,20}. The presence of the leader segments in the RNA species which protect fragments of sizes 520 and 440 nucleotides implies that the common tripartite leader is spliced onto their 5' end and that they originate from the major late transcription unit^{6,7}. When the HindIII-B fragment was protected with the selected RNA, digested with S₁, the ³²P-labelled BamHI-HindIII probe hybridized to a 520-nucleotide band, but not to one of 440 nucleotides (data not shown). This demonstrates that the body of the RNA species extends 520 nucleotides into the HindIII-B fragment, and that the 440-nucleotide fragment is encoded elsewhere on the genome (see below).

Recognition signals for splicing coincide with sites for termination of RNA polymerase III

To map the leader-body splice junction at the nucleotide level, we used a modification of the original Berk and Sharp S₁

method¹⁷. The ³²P-labelled *Hind*III-B fragment was cleaved with a restriction endonuclease, before hybridization to RNA. If a complementary RNA molecule extends across the restriction enzyme cleavage site, subsequent *S*₁ analysis will reveal two protected DNA fragments, which will add up to the length of the protected fragment obtained without cleavage.

Cytoplasmic RNA selected by hybridization to the *Bam*HI-*Hind*III fragment was used for this analysis. Figure 1C shows the cleavage sites for restriction enzymes used to cleave the *Hind*III-B fragment. Cleavage with endonuclease *Pvu*II resulted in the disappearance of the 520-nucleotide band, and the appearance of two new bands, 275 and 245 nucleotides long (Fig. 2B, lane b). The larger band represents the distance between the *Pvu*II and *Hind*III site at coordinate 31.5. Therefore, the RNA body extends 245 nucleotides to the left of the *Pvu*II site, positioning the leader-body splice site at coordinate 30.0. Cleavage of the DNA fragment with endonuclease *Sac*II (Fig. 2B, lane c) generated three bands of 125 (coordinates

Table 1 Early transcription of regions L1, L2 and L3 of the Ad2 genome

Coordinates DNA probe	Region	Length (kilobases)	c.p.m. in hybrid	Molarity
31.5-37.3	L1	2.0	240	1.0
41.0-50.1	L2	3.2	58	0.15
52.6-59.6	L3	1.5	187	1.0

Ad2-infected HeLa cells (500 FFU per cell) were resuspended at 5×10^7 cells per ml at 4 h.p.i. The cells were labelled with 500 μ Ci ml⁻¹ of ³H-uridine for 10 min at 37 °C, collected on frozen phosphate-buffered saline and nuclear RNA prepared³³. Nuclear RNA from an equivalent amount of cells was hybridized to various Ad2 DNA restriction fragments cloned in the bacterial plasmid pBR322. The cloned fragments were adsorbed to nitrocellulose filters and hybridization was carried out at 65 °C in 6 $\times$ SSC buffer (1 $\times$ SSC is 0.15 M NaCl, 15 mM sodium citrate). After hybridization the filters were washed extensively in 2 $\times$ SSC buffer and treated with pancreatic ribonuclease (20 μ g ml⁻¹) for 1 h at 37 °C. Hybridized RNA resistant to the nuclease treatment was measured by liquid scintillation counting. Hybridization to filters containing the pBR322 plasmid were used as background hybridization. This was in all cases fewer than 20 c.p.m. and has been subtracted from each value. The results are the mean from two independent experiments. The length in kilobases was calculated from the size of the Ad2 restriction fragments, assuming that the whole fragments were transcribed. Molarity was calculated by dividing the c.p.m. in hybrid by the length of each fragment and normalizing to 1.0 for the L1 probe (coordinates 31.5-37.3).

30.83-31.5), 235 (30.47-30.83) and 118 nucleotides (splice junction to 30.37), confirming that the splice junction is located at coordinate 30.0. The 5'-end splice site was confirmed by *S*₁ analysis using a 5'-end-labelled DNA probe (Fig. 2C). The band corresponding to the entire DNA probe is derived from RNA species *a* and *b* as they both extend leftward of the *Bam*HI cleavage site at coordinate 29.0.

The genes for the VA RNAs are located close to coordinate 30 on the viral genome^{8,9}. An artefactual cutting site may therefore be created during the *S*₁ analysis through displacement of the RNA body by VA RNA hybridizing to the same DNA template. To exclude this possibility, RNA was selected on fragment *Hind*III-J (coordinates 37.3-41.0), which is located outside the genes for the VA RNAs. *S*₁ analysis of this RNA generated the same pattern of bands as did RNA selected on fragment *Bam*HI-*Hind*III, but the bands corresponding to the VA RNAs are missing (Fig. 2B, lane f). This demonstrates that the splice site is not due to VA RNAs hybridizing to the DNA sequences representing the splice junction.

Because almost all known splice sites follow the G-T-A-G rule²¹, and because the nucleotide sequence in this region only contains one A-G dinucleotide, the site at which the third leader segment is spliced on to the RNA body is likely to be located at nucleotide 361 in Fig. 1D. The splice site for this RNA species is thus separated from the termination signal for VA RNA transcription⁹ by only 11 nucleotides (Fig. 1D). This compressed organization of processing signals suggests that the T-cluster, terminating transcription of VA RNA_{II}, also serves as a recognition signal for the enzyme(s) that cleave and ligate the tripartite leader to the RNA body.

Using blot hybridization to analyse DNA protected by *Bam*HI-*Hind*III-selected RNA from *S*₁ digestion, the total size of the RNA body was estimated to be 2,900 nucleotides (not shown). From its size and genomic location we conclude that this RNA corresponds to RNA species *c* (coordinates 30.0-38.5) from region L1 (Fig. 1B).

Identification of an additional leader segment on region L1 RNAs

Electron microscopy of RNA-DNA hybrids has shown that a fraction of late cytoplasmic RNAs from the L1 co-termination family contains an extra leader segment from coordinates 22.0 to 23.2. The 440-nucleotide band observed by *S*₁ analysis may correspond to this extra leader segment.

The origin of this band was studied by the *S*₁ endonuclease assay using RNAs selected by hybridization to the *Bam*HI-*Hind*III fragment. The ³²P-labelled *Hind*III-B fragment was cleaved with endonuclease *Xho*I before hybridization to the RNA and subsequent *S*₁ endonuclease treatment. The 440-nucleotide band was absent after this treatment and instead two

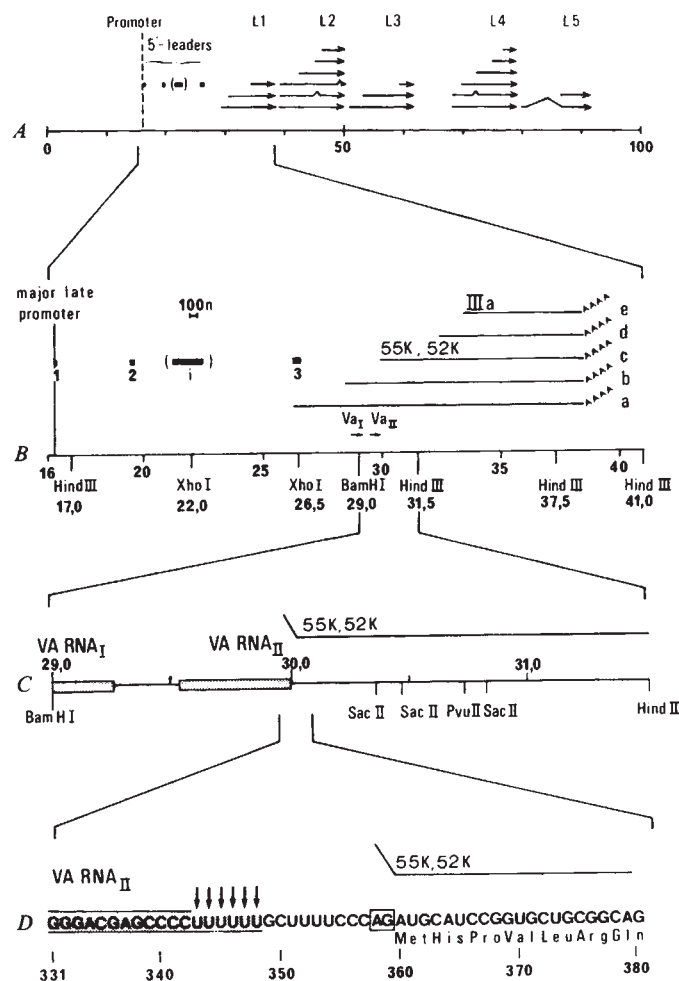


Fig. 1 Organization of the adenovirus 2 genome. **A**, Schematic diagram of RNAs originating from the major late transcription unit. The 20 identified RNA species are divided into five families (L1-L5), the members of each family having a common position for the poly(A) addition sites. The 5' leaders are spliced onto each of the 20 RNA bodies. The *i* leader found on a fraction of the RNAs is shown within brackets. **B**, Enlargement of region 16.0-41.0 encoding the major late promoter site at map coordinate 16.3. The common 5' leaders and the L1 3'-co-terminal family. The position of the VA RNA genes is shown as well as our revised RNA map for region L1. The position of restriction endonuclease cleavage sites relevant to this study are also shown. The mRNAs to which we have assigned translation products are indicated. **C**, Enlargement of fragment *Bam*HI-*Hind*III encoding the 5' sequences of the 55, 52K mRNA body. The position of restriction endonuclease cleavage sites used in this study are indicated. **D**, The leader-body splice junction of the 55, 52K mRNA. The nucleotide sequence is taken from ref. 9 and the numbering is from the *Bam*HI cleavage site at position 29.0. Arrows indicate the position of the heterogeneous 3' end of VA RNA_{II} and the box shows the A-G dinucleotide invariably found at the 3' side of intervening sequences²¹. The predicted amino-terminal sequence of the 55K and 52K polypeptides is also shown.

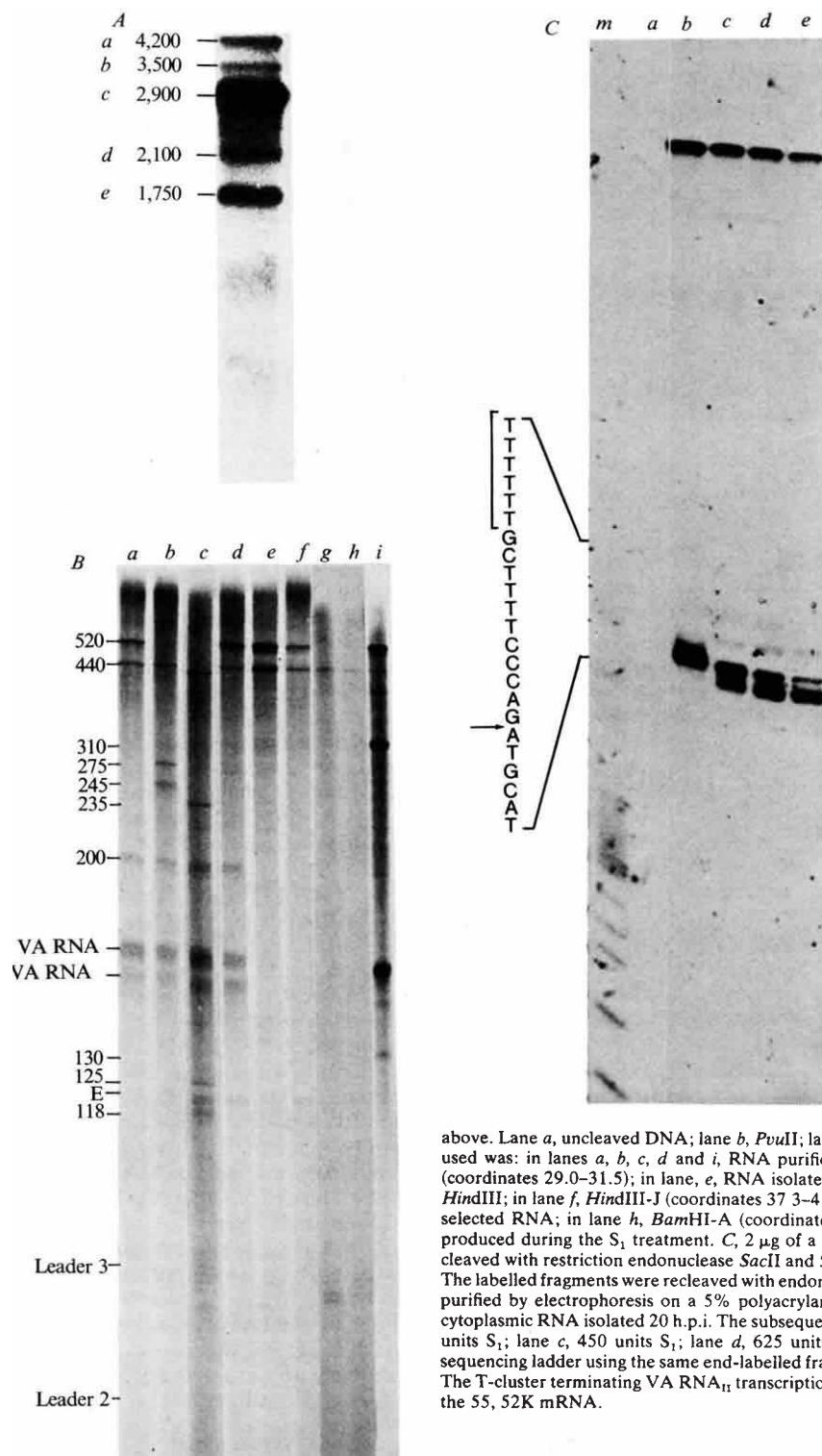


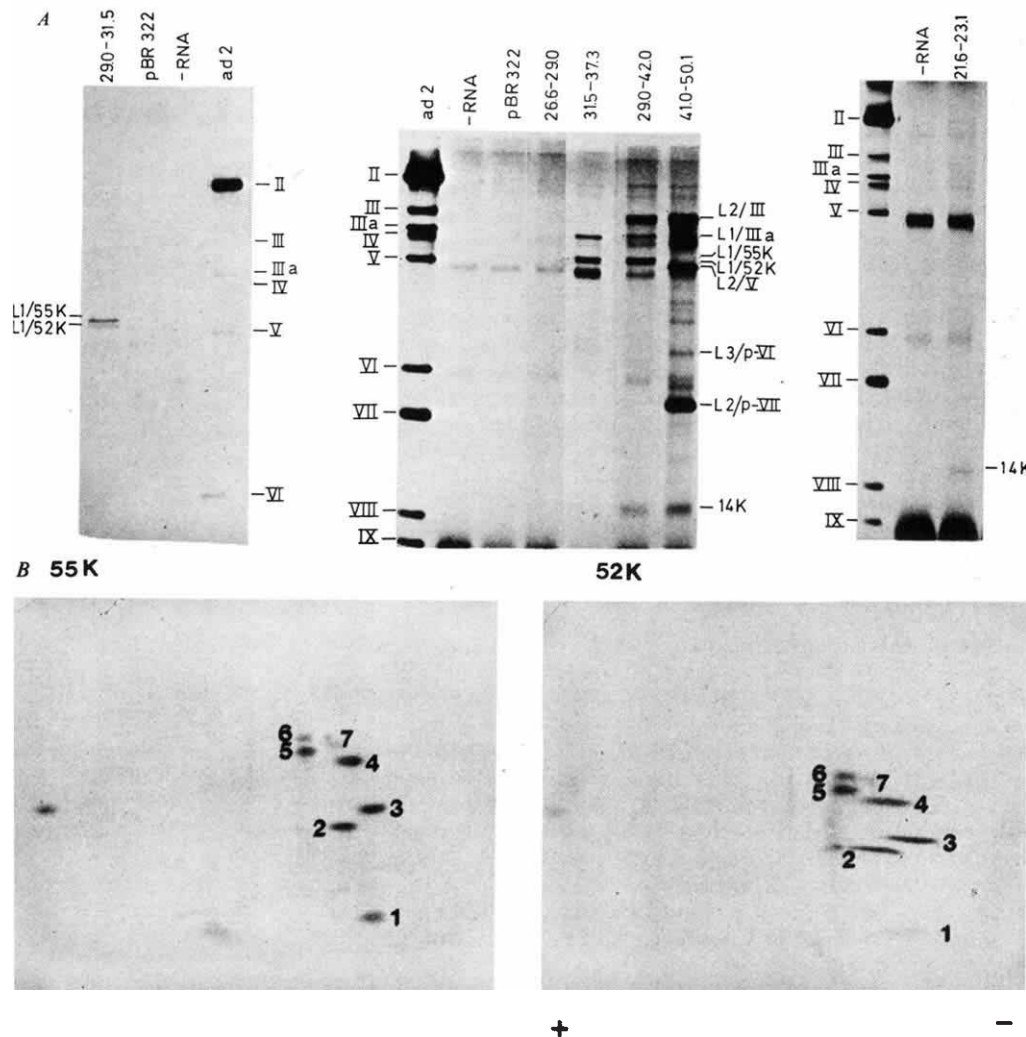
Fig. 2 S_1 endonuclease mapping of late adenovirus cytoplasmic RNA. **A**, 10 μ g of late cytoplasmic RNA, 10 μ g of tRNA and 2 μ g of Ad2 DNA were mixed, precipitated and dissolved in 10 μ l of 50 mM PIPES pH 6.4, 0.4 M NaCl, 1 mM EDTA and 80% formamide (buffer A). The DNA template was denatured by incubating the mixture at 70 °C for 10 min, followed by a hybridization for 3 h at 57 °C. The hybridization was terminated by the addition of 400 μ l of an ice-cold buffer containing 30 mM NaAc pH 4.4, 4 mM ZnSO₄, 0.28 M NaCl, 5 μ g of single-stranded calf thymus DNA and 5% glycerol. Then 225 units of endonuclease S_1 (Sigma) were added and the reaction mixture further incubated for 20 min at 37 °C. The S_1 cleavage was terminated by the addition of 50 μ l S_1 stop solution (3 M NaAc, 80 mM EDTA and 10 μ g tRNA) and 800 μ l ethanol. The precipitated material was collected by centrifugation and dissolved in 20 μ l of a buffer containing 50 mM NaOH, 1 mM NaOH, 1 mM EDTA, 2% Ficoll and marker dyes, and separated electrophoretically on a 2% alkaline agarose gel³⁴. After electrophoresis, the gel was soaked in 1 M Tris-HCl pH 7.5, 0.6 M NaCl for 10 min at room temperature and the S_1 -protected fragments transferred to a nitrocellulose sheet as described by Southern³⁵. The nitrocellulose filter was subsequently hybridized in 6 $\times$ SSC, 3 $\times$ Denhardt and 0.5% SDS for 12 h at 65 °C with a ³²P-labelled recombinant plasmid containing an insert of the HindIII-I fragment (coordinates 31.5–37.3) of Ad2 DNA¹⁶. The filters were washed in 0.2 $\times$ SSC, 0.5% SDS at room temperature for several hours, followed by a 15-min incubation at 65 °C and additional washes at room temperature. The blots were analysed by autoradiography at –70 °C using a DuPont intensifying screen. The HindIII fragments of Ad2 DNA run in a parallel slot on the agarose gel were used as size markers. **a–e** Relates the S_1 protected DNA fragments to the L1 mRNAs. **B**, Cytoplasmic RNA (as specified below) was mixed with 0.04 μ g of a recombinant plasmid containing the HindIII-B fragment (coordinates 17.0–31.5) of Ad2 DNA¹⁶, ³²P-labelled by nick translation. The mixture was ethanol precipitated and dissolved in 10 μ l of buffer A. The subsequent denaturation, hybridization and S_1 endonuclease treatment were as described above. The precipitate was dissolved in 10 μ l of a buffer containing 80% formamide, 1 mM EDTA and marker dyes. Electrophoretic separation was on thin 8% polyacrylamide gels containing 7 M urea and the ³²P-labelled bands were visualized by autoradiography. As size markers, ³²P-labelled Φ X174 DNA, cleaved with endonucleases HaeIII and HindII, was included in the gel. In lanes **b**, **c**, **d** and **i** the HindIII-B clone was cleaved with restriction endonucleases before S_1 treatment. The DNA fragments were mixed with RNA and processed as described

above. Lane **a**, uncleaved DNA; lane **b**, *PvuII*; lane **c**, *SacII*; lane **d**, *Sall*; lane **e**, *XhoI*-cleaved DNA. The RNA used was: in lanes **a**, **b**, **c**, **d** and **i**, RNA purified by hybridization selection³⁶ on fragment *BamHI*–HindIII (coordinates 29.0–31.5); in lane **e**, RNA isolated from ara C-treated cells and selected on fragment *BamHI*–HindIII; in lane **f**, HindIII-J (coordinates 37.3–41.0) selected RNA; in lane **g**, *BamHI*-C (coordinates 42–59.0) selected RNA; in lane **h**, *BamHI*-A (coordinates 59.0–100) selected RNA. **E** denotes an endogenous band produced during the S_1 treatment. **C**, 2 μ g of a recombinant plasmid containing the HindIII-B fragment was cleaved with restriction endonuclease *SacII* and 5'-end labelled using [γ -³²P]ATP and polynucleotide kinase³⁷. The labelled fragments were recleaved with endonuclease *BamHI* and the 480-base pair *BamHI*–*SacII* fragment purified by electrophoresis on a 5% polyacrylamide gel. The ³²P-labelled fragment was mixed with 3 μ g of cytoplasmic RNA isolated 20 h.p.i. The subsequent S_1 analysis was described above. Lane **a**, –RNA; lane **b**, 225 units S_1 ; lane **c**, 450 units S_1 ; lane **d**, 625 units S_1 ; lane **e**, 900 units S_1 ; lane **m** shows the purine-specific sequencing ladder using the same end-labelled fragment. The nucleotide sequence of the relevant part is shown. The T-cluster terminating VA RNA_{II} transcription is shown in brackets and the arrow indicates the splice site of the 55, 52K mRNA.

new bands of 310 and 130 nucleotides were detected (Fig. 2B, lane **i**). A similar analysis using endonuclease *Sall* (cleavage sites 26.0 and 26.6) did not cleave the 440-nucleotide band (Fig. 2B, lane **d**). This shows that the 440-nucleotide band is derived from sequences located around the *XhoI* cleavage site at coordinate 22.0 (Fig. 1B). The size and location of the 440-nucleotide band suggest that it is identical to the extra leader, designated the *i* leader, previously identified by electron microscopy¹². As the RNAs used in the S_1 analysis were selected on DNA fragments outside the coding region for the *i* leader, the results suggest that it is covalently attached to the RNA molecules. RNA selected on fragments *BamHI*-C (coordinates 41.0–

59.0) and *BamHI*-A (coordinates 59.0–100) also generated the 440-nucleotide band after S_1 analysis (Fig. 2B, lanes **g**, **h**). This demonstrates that the *i* leader is present on at least a fraction of RNAs derived from all regions of the major late transcription unit. Densitometer tracing of the 520- and 440-nucleotide bands (Fig. 2B, lane **a**) indicates that there are two types of L1 RNAs, differing with respect to their 5'-leader segments. Type I RNAs have the structure 'leader 1, 2, 3 RNA body', and comprise ~40% of the total population, while type II mRNAs have the structure 'leader 1, 2, *i*, 3, RNA body' and comprise ~60% of the total. The 520-nucleotide long band and the leader fragments 2 and 3 were all present in equimolar concentrations.

Fig. 3 *In vitro* translation of hybridization selected mRNAs. **A**, Late cytoplasmic Ad2 RNA³⁸ was purified by hybridization to restriction fragments of Ad2 DNA immobilized on nitrocellulose filters³⁶. The hybridization-selected RNAs were translated in a reticulocyte cell-free system and the translation products were analysed by SDS-polyacrylamide gel electrophoresis³⁸ (SDS-PAGE). The gels were analysed by fluorography. The *Hind*III restriction fragments were cloned in the plasmid pBR322¹⁶ and the hybrid plasmids were used for selection. Hybridization to the bacterial plasmid, pBR322; was used as a control for the specificity of selection. The endogenous synthesis in the cell-free system is shown in slots indicated by -RNA. Ad2 denotes here and in subsequent figures a ³⁵S-methionine-labelled Ad2 marker virus. Coordinates for restriction fragments used for hybridization selection are shown at the top of the figure. Restriction fragments used were as follows: *Sac*I-M and -N (21.6–23.1), *Sal*I-*Bam*HI (26.6–29.0), *Bam*HI-*Hind*III (29.0–31.5), *Hind*III-I (31.5–37.3), *Bam*HI-D (29.0–42.0), *Hind*III-D (41.0–50.1). The left panel represents a 10% polyacrylamide gel and the right panels 13% polyacrylamide gels. **B**, Tryptic fingerprint analysis of the 55K and 52K polypeptides. Messenger RNA selected by hybridization to fragment *Bam*HI-*Hind*III (coordinates 29.0–31.5) was translated in a reticulocyte cell-free system in the presence of ³⁵S-methionine. The translation products were separated on a 10% SDS-polyacrylamide gel and the 55K and 52K polypeptides were excised from the gel. Eluted polypeptides were oxidized with performic acid and digested with TPCK-treated trypsin³⁹. Tryptic peptides were separated by electrophoresis in the first dimension followed by ascending chromatography in the second dimension. ³⁵S-methionine-labelled tryptic peptides were visualized by autoradiography.



Polypeptides encoded in region L1

To test whether the RNA species encoded in region L1 code for polypeptides, late cytoplasmic RNA selected on the *Bam*HI-*Hind*III fragment was used to programme an *in vitro* translation system. The selected RNA directed the synthesis of two polypeptides with molecular weights of 55,000 (55K) and 52,000 (52K). In addition, the third member of the L1 3'-co-terminal family, polypeptide IIIa²², was detected when RNA was selected by hybridization to fragment *Hind*III-I (coordinates 31.7–37.3) (Fig. 3A). RNA selected on fragment *Bam*HI-D (coordinates 29.0–42.0) directed the synthesis of the 55K, 52K and IIIa polypeptides from region L1 and polypeptide III from region L2 (Fig. 3A), while the L2 polypeptides III, pVII and V were detected with RNA selected on fragment *Hind*III-D (coordinates 41.0–50.1). RNA selected on fragment *Sal*I-*Bam*HI (coordinates 26.6–29.0) did not yield any polypeptides (Fig. 3A). This region is absent in all the L1 RNA molecules except in species *a* and *b* in Fig. 1B. Interestingly, a 14K polypeptide was synthesized from RNAs selected by hybridization to all DNA probes covering regions L2–L5 (Fig. 3A).

Together, these results demonstrate that the 55K and 52K polypeptides are encoded by mRNA species *c*, mapping between coordinates 30.0 and 38.5 (Fig. 1B), and that polypeptide IIIa is encoded by mRNA species *d* or *e* (coordinates 33.5–38.5). We tentatively assign polypeptide IIIa to mRNA species *e* (coordinates 33.5–38.5) based on the enhanced accumulation of this species late in infection (see below). RNA selected on the viral r-strand directed the synthesis of the 55K

and 52K polypeptides, confirming that the mRNA is transcribed in a rightward direction (not shown).

The structural relationship between the 55K and the 52K polypeptides was investigated by tryptic fingerprint analysis. A very similar but not identical set of ³⁵S-labelled peptides was obtained (Fig. 3B), suggesting that these two polypeptides are structurally related. They may therefore be encoded on the same mRNA and their size difference may reflect a post-translational modification which occurs *in vitro*. We cannot, however, exclude the possibility that the 52K polypeptide is encoded by a minor fraction of species *c* which has an internal splice in the mRNA body.

Presence of the *i* leader alters the translation of the mRNA

The 55, 52K mRNA was selected by hybridization to fragments *Sac*I-M and -N (coordinates 21.6–23.1) which cover the region of the genome encoding the *i* leader (not shown). This result is expected because around 60% of the 55, 52K mRNA has the structure 'leader 1, 2, *i*, 3, RNA body'. However, cell-free translation with *Sac*I-M- and -N-selected mRNA did not result in the synthesis of the 55K and 52K polypeptides (Fig. 3A), but instead, produced a 14K polypeptide; this suggests that mRNA species containing the *i* leader do not translate the body of their mRNA sequences. Because eukaryotic ribosomes usually initiate translation at the AUG codon closest to the 5'-end²³ of the mRNA, the results suggest that an AUG triplet is encoded within the *i* leader itself. Late cytoplasmic RNA selected on

from regions L2 and L3 were not detected in the absence of virus DNA replication (Fig. 5A).

The selective accumulation of mRNA species *c* early in infection was also demonstrated by *in vitro* translation of RNA selected on fragment HindIII-J (coordinates 37.3–41.0). RNA isolated from cells treated with ara C yielded the 55K and 52K polypeptides and the 14K polypeptide, whereas cytoplasmic RNA from untreated cells obtained late in infection directed the synthesis of polypeptides III, IIIa, 55K and 52K (Fig. 5B).

Role of newly synthesized viral proteins in major late adenovirus promoter activation

To determine whether virus-specified early proteins are necessary for activation of the major late promoter, we analysed RNA samples prepared from cells maintained in conditions of stringent inhibition of protein synthesis²⁴. Treatment with puromycin or anisomycin from the start of the infection still resulted in selective accumulation of the 55, 52K mRNA (Fig. 5A, slots c, d). Furthermore, the 55, 52K mRNA synthesized in the absence of virus protein synthesis was functionally active when assayed in an *in vitro* translation system (Fig. 5B).

The selective expression of the 55, 52K mRNA from the major late transcription unit, in the absence of protein synthesis, suggests that the activation of the major late promoter, and the splicing and polyadenylation of the 55, 52K mRNA, do not require newly synthesized viral polypeptides. After completion of this study, Lewis and Mathews²⁵ reported similar results concerning the synthesis of the 55K and 52K polypeptides.

Truncated nuclear RNAs accumulate in the absence of virus DNA replication

When virus DNA replication is blocked, cytoplasmic RNAs from region L2 and L3 are undetectable. To determine if the nuclear precursors of these RNAs were present we examined nuclear RNA preparations.

On *S*₁ analysis of nuclear RNA prepared from cells 18 h.p.i. and not treated with metabolic inhibitors, all cytoplasmic RNA species specific for regions L1, L2 and L3 were detected in nuclear RNA preparations, although the relative abundances in the nucleus differed from those in the cytoplasm. Nuclear

accumulation seems to be greatest for the longest RNAs from each 3'-co-terminal family. The most prominent nuclear RNAs accumulated from region L1 were species *a* and *c* (Fig. 6, slot *a*), but a cluster of long RNAs, ranging in size from 6,000 to 8,000 nucleotides, was also detected, probably representing precursor molecules originating from the major late transcription unit²⁶. Large nuclear RNA species also accumulated from region L2 (bands of 4,400, 3,800 and 3,300 nucleotides, Fig. 6, slot *e*). Region L3 accumulated nuclear RNAs corresponding to the 4,400- and 3,600-nucleotide long pVI and hexon mRNAs²⁷ (Fig. 6, slot *g*). The 4,400-nucleotide band seen with the L2 and L3 probes corresponds to the same RNA species, as the pVI mRNA from region L3 extends ~300 nucleotides into the HindIII-D fragment²⁷.

Analysis of nuclear RNAs, prepared from cells where virus DNA replication was blocked, showed a selective accumulation of RNA species *a* and *c* from region L1 (Fig. 6, slots *b–d*). We also detected RNA molecules co-linear with the DNA template from the promoter site at coordinate 16.3 up to the poly(A) addition site at coordinate 38.5. Similar results were obtained when protein synthesis was inhibited with puromycin or anisomycin from the start of virus infection.

Nuclear RNA species from region L2 were barely detectable in RNA preparations from cells treated with ara C (Fig. 6, slot *f*) and only the 3,300-nucleotide long RNA species was discerned. Nuclear RNA from region L3 was not detected (Fig. 6, slot *h*), demonstrating that, in the absence of virus DNA replication, RNA beyond coordinate 59 does not accumulate.

To determine the length of the transcript from the major late promoter at early times after infection, cells were pulse-labelled for 10 min at 4 h.p.i. with ³H-uridine and nuclear RNA was prepared. The nuclear RNA was hybridized to cloned DNA fragments covering different regions of the major late transcription unit. Table 1 shows that transcripts from region L1 were 5–10-fold more abundant than from region L2 early in infection. However, there was marked hybridization to the probe from region L3, probably reflecting the fact that the leftward transcript for early region 2 terminates within the double-stranded DNA probe used for the L3 region (cited in ref. 28). The low hybridization to region L2 compared with region L1 together with the *S*₁ analysis of nuclear RNA suggest that transcription terminates within the L2 region at early times of infection. Similar results were obtained with pulse-labelled RNA from cells treated with anisomycin from the start of virus infection (not shown). The length of the transcript from the major late promoter early in infection was independently determined by Shaw and Ziff²⁹.

Conclusions

We have analysed RNAs originating from the major late transcription units of adenovirus type 2, at different times after infection. The results suggest that the shift from the early to the late phase of virus replication does not involve activation of the major late promoter at map coordinate 16.3. Rather, this promoter is already active during the early phase. Furthermore, the synthesis of viral polypeptides is not required for its activation, as mRNA from the major late promoter is produced under stringent inhibition of protein synthesis²⁴. A preferential synthesis of mRNA species *c* from region L1 (Fig. 1B), which encodes the two structurally related polypeptides of molecular weight 55,000 and 52,000, was observed at early times. The lack of detectable amounts of L2 and L3 mRNAs may be explained by the fact that nuclear transcripts produced at early times of infection seem to terminate within the L2 region. Therefore, the shift from the early to the late phase of virus replication may involve a mechanism which allows the production of larger transcripts from the major late transcription unit. This mechanism could be mediated by virus-coded factors which associate with the RNA polymerase, by a change in the nucleoprotein structure of the DNA template, or by DNA replication providing single strands of virus DNA which may serve as a more efficient template for late transcription.

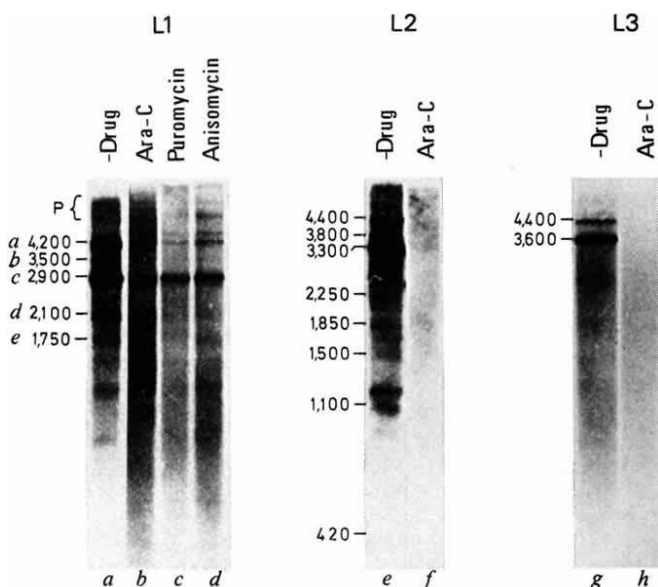


Fig. 6 *S*₁ blot analysis of nuclear RNA prepared from cells maintained in the presence of metabolic inhibitors. 10 µg of nuclear RNA, 10 µg of tRNA and 2 µg of Ad2 DNA were mixed, precipitated and analysed by the *S*₁ blot technique described in Fig. 2 legend. The nuclear RNA was prepared by the method described by Smith *et al.*³³ from cells maintained in the presence of ara C (25 µg ml⁻¹; lanes *b, f, h*) from 1 to 16 h.p.i. and puromycin (100 µM; lane *c*) and anisomycin (100 µM; lane *d*) from the onset of infection to 7 h.p.i. Nuclear RNA prepared 16 h.p.i. from cells grown without any metabolic inhibitors was also used (lanes *a, e, g*). Hybridization was with the L1-, L2- and L3-specific probes described in Fig. 5 legend.

The selective accumulation of the 55, 52K mRNA early in infection may be explained by a virus-regulated abundance control of late mRNA production. At different times after infection the individual mRNAs from region L1 (Fig. 1B, species *a-e*) are differentially expressed. The first mRNA to accumulate in the cytoplasm is the 55, 52K mRNA, whereas at late times of infection, species *e*, encoding virion polypeptide IIIa, is the most abundant. Thus, the selective appearance of the 55, 52K mRNA early in infection may reflect a regulation of the abundance of region L1 mRNAs. This regulation is probably accomplished by virus-regulated changes in the splicing pattern of the transcript but may also depend on virus-regulated changes in cytoplasmic mRNA stabilities. However, the latter alternative is unlikely because the mRNA stability seems to be enhanced at late times of infection³⁰.

As RNA polymerase III transcription normally terminates within a stretch of U residues³¹, and the termination signal for VA RNA_{II} transcription is separated by only 11 nucleotides from the splice junction of the 55, 52K mRNA (Fig. 1D), the

acceptor site of the 55, 52K mRNA has a U-rich sequence directly upstream from its splice junction. This sequence arrangement resembles splice sites found in several eukaryotic genes³². If the abundance control of late mRNA production involves virus-controlled splicing events, the appearance of the 55, 52K mRNA at early times of infection may reflect a preference for the acceptor site of the 55, 52K mRNA.

The 55, 52K mRNA exists at about equal frequency in two forms which differ by the presence of the *i* leader. In the RNA species which have an insertion of the *i* leader within the tripartite leader, the expression of the polypeptides from the mRNA body seems to be blocked. This suggests that insertion of the *i* leader, by splicing, controls gene expression at the level of translation.

We thank Ms M. Pettersson for technical assistance, Dr M. Perriacaudet for the gift of the cDNA clones, Drs L. Philipson, U. Pettersson and V. Tate for help in the preparation of the manuscript and Ms M. Gustafson and C. Sjöholm for secretarial help. This work was supported by grants from the Swedish Society against Cancer and the Swedish MRC.

Received 22 December 1980; accepted 11 May 1981.

- Ziff, E. B. & Evans, R. M. *Cell* **15**, 1463-1475 (1978).
- Evans, R. M. *et al.* *Cell* **12**, 733-739 (1977).
- Nevins, J. R. & Darnell, J. E. *J. Virol.* **25**, 811-823 (1978).
- Ziff, E. B. & Fraser, N. J. *J. Virol.* **25**, 897-906 (1978).
- McGrogan, M. & Raskas, H.-J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 625-629 (1978).
- Berget, S. M., Moore, C. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3171-3175 (1977).
- Chow, L. T., Gelin, R., Broker, T. & Roberts, R. J. *Cell* **12**, 1-8 (1977).
- Mathews, M. B. & Pettersson, U. *J. molec. Biol.* **119**, 293-328 (1978).
- Akusjärvi, G., Mathews, M. B., Andersson, P., Vennström, B. & Pettersson, U. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2424-2428 (1980).
- Chow, L. T. & Broker, T. *Cell* **15**, 497-510 (1978).
- Fraser, N. W., Sehgal, P. B. & Darnell, J. E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2571-2575 (1979).
- Chow, L. T., Broker, T. R. & Lewis, J. B. *J. molec. Biol.* **134**, 265-303 (1979).
- Kitchingman, G. R. & Westphal, H. J. *J. molec. Biol.* **137**, 23-48 (1980).
- Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. *Cell* **18**, 469-484 (1979).
- Waslyk, B., Kedinger, C., Corden, J., Brison, O. & Chambon, P. *Nature* **285**, 367-372 (1980).
- Stenlund, A., Perriacaudet, M., Tiollais, P. & Pettersson, U. *Gene* **10**, 47-52 (1980).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 45-56 (1977).
- Weinmann, R., Brendler, T. G., Raskas, H. J. & Roeder, R. G. *Cell* **7**, 557-566 (1976).
- Akusjärvi, G. & Pettersson, U. *J. molec. Biol.* **134**, 143-158 (1979).
- Zain, S., Giergas, T. R., Bullock, P., Wong, G. & Gelinas, R. E. *J. molec. Biol.* **135**, 413-433 (1979).
- Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853-4857 (1978).
- Lewis, J. B., Anderson, C. W. & Atkins, J. F. *Cell* **12**, 37-44 (1977).
- Kozak, M. *Cell* **15**, 1109-1123 (1978).
- Persson, H., Monstein, H.-J., Akusjärvi, G. & Philipson, L. *Cell* **23**, 485-496 (1981).
- Lewis, J. B. & Mathews, M. B. *Cell* **21**, 303-313 (1980).
- Berget, S. M. & Sharp, P. A. *J. molec. Biol.* **129**, 547-565 (1979).
- Akusjärvi, G. & Persson, H. *J. Virol.* **38**, 469-482 (1981).
- Nevins, J. R. *et al.* *J. Virol.* **32**, 727-733 (1979).
- Shaw, A. R. & Ziff, E. B. *Cell* **22**, 905-916 (1980).
- Wilson, M. W., Nevins, J. & Darnell, J. E. *Cold Spring Harb. Symp. quant. Biol.* **44**, 447-455 (1979).
- Korn, L. J. & Brown, D. D. *Cell* **15**, 1145-1156 (1978).
- Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. *Nucleic Acids Res.* **8**, 127-142 (1980).
- Smith, M. M., Reeve, A. E. & Huang, R. C. C. *Cell* **15**, 615-626 (1978).
- Favaloro, J., Treisman, R. & Kamen, R. *Meth. Enzym.* (in the press).
- Southern, E. M. *J. molec. Biol.* **98**, 503-518 (1975).
- McGrogan, M., Spector, D. J., Goldenberg, C. J., Halbert, D. & Raskas, H. J. *Nucleic Acids Res.* **6**, 593-607 (1979).
- Weaver, R. S. & Weissmann, C. W. *Nucleic Acids Res.* **7**, 1175-1193 (1979).
- Persson, H., Pettersson, U. & Mathews, M. B. *Virology* **90**, 67-79 (1978).
- Persson, H., Jansson, M. & Philipson, L. *J. molec. Biol.* **137**, 375-394 (1980).

Diversity of germ-line immunoglobulin V_H genes

David Givol, Rina Zakut, Kim Effron, Gideon Rechavi,
Daniela Ram & Justus B. Cohen

Department of Chemical Immunology, The Weizmann Institute of Science, Rehovot 76100, Israel

The sequences of four embryonic mouse immunoglobulin V_H genes have been compared. All genes end at codon 98 and code for a hydrophobic signal peptide of 19 residues interrupted at codon -4 by an intron of 83 base pairs. Substitutions occur in all gene segments but at a significantly higher frequency in the hypervariable regions. The data suggest an evolutionary basis for the diversity of immunoglobulin genes. Divergence resulted also in a termination codon in two of the genes, suggesting that part of the V gene repertoire cannot be expressed unless some correction mechanism is available.

IMMUNOGLOBULINS are composed of heavy and light chains. Their variable region (V_L and V_H) is confined to the amino-terminal domain of the molecule and comprises the antibody-combining site. The combining site is formed by three hypervariable segments of the V region which contain the complementarity-determining residues (CDR1-3) for antigen^{1,2}. The variable region of the light chain is encoded in the germ line by three separate gene segments: the leader (L), coding for a hydrophobic signal peptide which is not present in the mature protein, the variable segment (V_L) and the joining segment (J_L)³⁻⁵. The variable region of the heavy chain is encoded by four gene segments: the leader (L), the variable segment (V_H), the 'diversity' segment (D) and the joining segment (J_H). The D and J_H segments contain the codons for CDR3 whereas the V_H segment contains the codons for CDR1 and CDR2 (refs 6, 7). It has been shown that there are only four J_H gene segments⁷. Amino acid sequence data indicated that the

genes for the variable regions are organized in subgroups which evolved by gene duplication^{8,9}. Expansion and contraction of these subgroups and evolutionary accumulation of mutations contribute to the diversity of these genes^{8,9}. Recent analysis at the DNA level demonstrated multiple genes in subgroups of both V_L¹⁰ and V_H¹¹.

A basic problem in immunology is the extent of inherited diversity compared with the diversity generated somatically in the V genes. The V_H segment is present in two different DNA environments during the life of the antibody-forming cells, the embryonic and the mature environments. During lymphocyte differentiation DNA rearrangement joins V with D and J (V-D-J joining) to form the active V_H gene⁶. As evidence for recombination of V_H with each of the four J_H segments has been found^{6,7} and the joining is flexible, part of the V-region diversity in CDR3 is generated somatically by this recombination. These recombination events involve the cutting and joining of DNA

Table 1 Per cent substitution in V_H -gene segments

Clone	5' region	Non-coding segments			Signal peptide		Coding segments			
		Intron	3' region		R	S	FR	S	R	CDR
pCh108B	15	15.7	20.0		7.0	1.8	4	4	11.6	0
pCh104	23.8	25.3	37.9		12.3	1.8	5.34	4.4	30.4	1.5
pCh111	23.8	21.7	37.9		10.5	5.3	6.2	5.8	27.5	4.3

The per cent substitution was calculated as compared with the pCh108A sequence. In the coding region substitutions are divided into replacement (R) and silent (S) substitutions. The 3' region was calculated beyond the recombination signals. FR, framework regions of V_H . CDR, complementarity-determining residue regions of V_H .

segments and may be accompanied by other enzymatic activities, such as error-prone repair^{12,13}, which result in further substitutions along the V_H segment. The question arises as to whether the hypervariability found in CDR1 and -2 pre-exists in the germ line or is generated in the new DNA environment only after somatic rearrangement.

To analyse this question we isolated and sequenced four mouse V_H genes isolated from a BALB/c mouse embryo DNA library¹⁴. These genes belong to the V_H II subgroup. Their nucleotide sequence shows 80–90% homology with the cDNA clone of mouse myeloma MPC11 heavy chain¹⁵ that was used as a probe in isolating these genes. The data indicate that the hypervariability in the regions corresponding to CDR1 and -2 is already present in the embryonic non-rearranged V_H genes, suggesting that the substitution in hypervariable regions is a result of an evolutionary process. The sequence also shows that two of the V_H genes contain termination codons within their coding region and may be considered pseudogenes. Perhaps this is the price paid for keeping a multiple gene family like the immunoglobulin V genes.

Characterization of V_H -containing clones

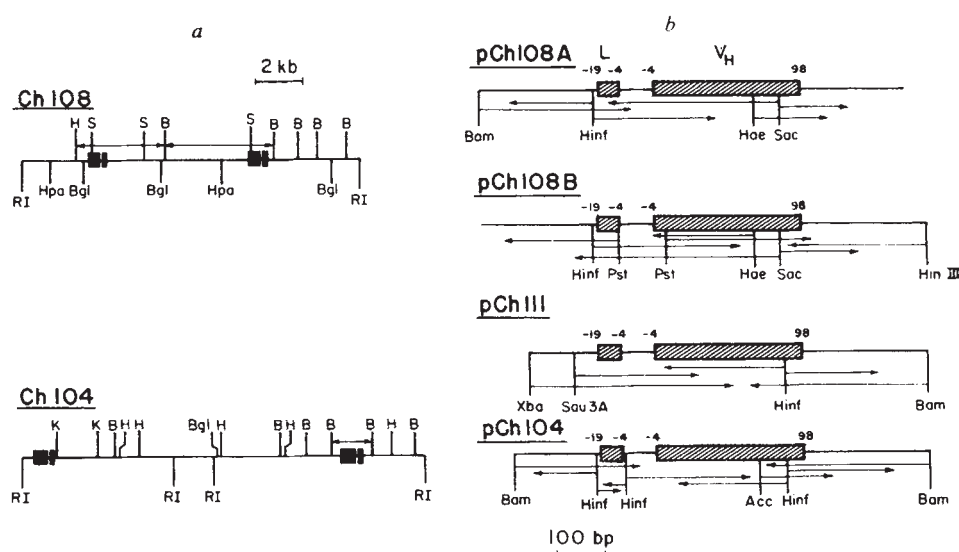
We isolated eight cross-hybridizing clones by screening our mouse embryo DNA library with plasmid pV(11)¹² which contains V_H cDNA derived from myeloma MPC11 (ref. 15). Three of these clones that showed strong cross-hybridization were selected for sequencing. One of them (Ch111) contained only one V_H gene; the others contained two V_H genes. In clone Ch108 the two V_H genes are located on a single *Eco*RI fragment with a spacing of ~7 kilobases (kb) (Fig. 1a) suggesting that V_H genes which belong to the same subgroup are close to one another in the genome. Clone Ch104 contained two V_H genes on

separate *Eco*RI fragments with a spacing of ~12 kb. The appropriate segments containing the V_H genes were subcloned in pBR322 and subjected to DNA sequencing, according to the strategy shown in Fig. 1b. The sequence of ~800 base pairs (bp) was determined for each of the genes, and compared with that of MPC11 cDNA to locate the coding region. The overall structure of all four genes demonstrates their germ-line nature as the coding sequence ends prematurely at codon 98 and contains neither D nor J segments. The codon corresponding to amino acid 1 is preceded by a sequence that codes for a signal peptide^{16,17} of 19 amino acids which is interrupted at codon -4 by an intron of 83 bp.

Sequence of the 5'-noncoding and leader region

Figure 2 depicts the sequence of the four mouse embryonic V_H genes upstream from codon 1 of the mature protein, in comparison with MPC11 cDNA sequence¹⁵. It is clear that the V_H genes, although homologous, are variable in the region corresponding in part to the 5'-untranslated region of the mRNA. Hence, the variability is not confined to the V region, but extends into the 5'-untranslated region. Thus heavy-chain mRNA has a variable 5'-untranslated region while the 3'-untranslated region is constant¹⁸. The sequence coding for the 19 residues which compose the hydrophobic signal polypeptide of the precursor heavy chain is separated at residue -4 (Gly) by an intron of 83 nucleotides which starts with GT and ends with AG. The size of this intron in all four genes is identical and differs by 2 nucleotides from the size of another mouse V_H intron⁷ and by 21 nucleotides from a human V_H intron which is 104 bp long¹⁹. It is significant that the location of this intron at codon -4 is identical in mouse V_{κ} ^{5,20}, V_{λ} ²¹ and V_H ^{6,7}, as well as in

Fig. 1 Restriction enzyme maps of the mouse DNA inserts in clones Ch108 and Ch104 (a), and strategy used in sequencing of the subclones pCh108A, pCh108B, pCh111 and pCh104 (b). Cleavage sites were determined by single and double digestion of DNA with various restriction enzymes and by the end-labelling method of Smith and Birnstiel³³. The DNA segments between arrows in a were subcloned in pBR322. Sequence determination was by the chemical degradation method³⁴ of end-labelled fragments. End-labelling was done at the 3' end by site filling of 5'-protruding ends with reverse transcriptase. The reaction mixture (20 μ l) contained 1–5 μ g of DNA fragments, 20 μ Ci of one or two [α -³²P]deoxynucleoside triphosphates and 5 units of reverse transcriptase, and was incubated at 37°C for 30 min. Blunt-ended fragments were labelled at the 3' end using DNA polymerase (n.Klenow) and the appropriate [α -³²P]dNTP. 3'-protruding ends were labelled with [α -³²P]cordycepin triphosphate and terminal deoxynucleotidyl transferase³⁵. Enzymes used: RI = *Eco*RI; B = *Bam*HI; H = *Hind*III; Bgl = *Bgl*II; K = *Kpn*I; S = *Sac*I; Hpa = *Hpa*II; Hae = *Hae*II; Acc = *Acc*I; Hinf = *Hinf*I; Pst = *Pst*I; Sau = *Sau*3A. In b the numbers above the schemes represent codons for amino acids.



--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Fig. 2 Sequence of the 5' segment of four embryonic V_H genes including the codons for the signal peptide. Identical nucleotides are represented by dashes, insertions are introduced above the line and an empty space indicates a deletion. The region between parentheses was not determined (ND). Amino acid residues predicted by the nucleotide sequence are given above the line of the codons. The insert of MPC11 (M11) cDNA clone¹⁵ begins at codon -14 and the coding sequence of the signal peptide was determined in this cDNA clone (unpublished work).

the human V_H gene¹⁹. The homology at the two ends of the introns (first 30 bp and last 25 bp) in our genes is much greater than that in the middle part. The sequence of another reported V_H intron⁷ is significantly different (50% substitution) from the four sequences shown here (18–25% substitution) and indicates that V_H genes of different subgroups have diverged markedly from one another. The coding sequence for the signal peptide is not identical within the subgroup and the per cent substitution is similar to that of the V_H -coding region. However, comparison with sequences coding for signal peptides in V_H genes of other subgroups^{6,7,22} shows a greater divergence from our genes. The accumulation of mutations is proportional to the divergence time of the genes, and our comparison indicates that the leader sequence is part of the V_H duplication unit. Note that the middle highly hydrophobic region of this predicted amino acid sequence (residues -15 to -3) is very similar in all four genes and changes are more significant at the two ends of the sequence coding for the signal peptide.

The data also show that clone pCh108B has a substitution which changes the first Met (ATG) into Ile (ATA). Unless there is some correction mechanism, the mRNA of this gene may not be translated into an active protein. The validity of this substitution was checked in the original Ch108 clone.

Sequence of V_H -coding region and the recombination signals

Figure 3 shows the sequence of the V_H -coding region and the sequence downstream (3') from it. Comparison with the MPC11 sequence¹⁵ and with known J_H segments²³ demonstrates that all four genes end prematurely at Arg 98 which is exactly the end of framework III segment (no. 94 in the numbering system of Kabat *et al.*²³). Hence the D segment in the MPC11 V region starts at the beginning of CDR3 and contains eight amino acids (Gly-Ile-Tyr-Tyr-Asn-Ser-Ser-Pro). The V_H -coding sequence clearly demonstrates that substitutions occur preferentially at CDR regions (see later for comparison). When substitutions at the framework regions do occur, they are usually confined to the same position in all four genes. Most of these positions, such as codons 13, 16, 40, 41, 43 and 74, code for amino acid residues in which there is little structural constraint on replacements in the folded V domain. Comparison of these positions with homologous residues in three-dimensional models of V regions demonstrates that these residues are present in positions which

are completely exposed to solvent. It was shown at the protein level that amino acids which are buried in the domain are more conserved than those that are exposed to solvent^{24,25}. Many of the substitutions in the framework result in two alternative codons causing conservative changes of amino acids. In several places the extent of diversity is not clear from this limited analysis. For example, codons 81–85 may be analogous to the fourth hypervariable region²⁶ of human V_H and the limited variations shown here may indicate the presence of this region in mouse embryonic genes.

At the 3' end of these V_H genes there are two recombination signals separated by 23 or 21 (pCh111) bp. The first signal

CACAGTG starts immediately after the codon for Arg 98 whereas in other cases there are 2 or 3 bp between V_H and this signal. This sequence is similar to that found in mouse V_H and V_L and in human V_H genes^{3,5-7,19} regardless of subgroups, and it is inversely complementary to a heptanucleotide present just 5' to the J_H region. As also noted by others, the second signal nonanucleotide is dA rich and contains more substitutions than the first signal. The sequence between the two signal oligonucleotides is highly homologous within this subgroup while the sequence of the 3' region beyond the second signal diverged more than other segments in these genes.

The immune V-gene repertoire may be rich in pseudogenes

Two of the genes, clones pCh104 and pCh111, contain termination codons at positions 39 and 94 respectively. Because there are some 40 substitutions in the V_H -coding region (294 bp long), a high proportion of these genes may have one substitution leading to a terminator. This is perhaps the price paid for keeping a multiple gene family (like immunoglobulin V genes) in evolution. The survival of the animal does not depend on any individual member of the V-gene family and selection operates on the entire V-gene repertoire to increase the number and diversity of V genes. The evolutionary accumulation of substitutions in V genes is essential for the generation of antibody diversity, and it is unavoidable that some of these genes will also acquire termination codons. The V genes which contain terminators may be pseudogenes²⁷. However, they differ from other pseudogenes reported. Recently Bentley and Rabbitts²⁸ described a human V_κ pseudogene containing terminators,

repertoire must be degenerate and contains nonsense mutations and pseudogenes. We may speculate that some correction processes, not yet found, have developed in lymphocytes to enable them to use the genetic potential encoded in such genes. The CDR regions, which are most essential for the function of the antibody, are hypervariable in the germ-line genes and after

Received 16 February; accepted 19 May 1981.

1. Wu, T. T. & Kabat, E. A. *J. exp. Med.* **132**, 211–250 (1970).
2. Kabat, E. A. & Wu, T. T. *Ann. N.Y. Acad. Sci.* **190**, 382–391 (1971).
3. Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15**, 1133–1144 (1978).
4. Brack, C., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15**, 1–14 (1978).
5. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450–3454 (1979).
6. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* **19**, 981–992 (1980).
7. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676–683 (1980).
8. Milstein, C. & Pink, J. R. L. *Prog. Biophys. molec. Biol.* **21**, 211–263 (1970).
9. Smith, G. P., Hood, L. & Fitch, W. M. A. *Rev. Biochem.* **40**, 969–1012 (1971).
10. Seidman, J. G., Leder, A., Nau, M., Norman, B. & Leder, P. *Science* **202**, 11–17 (1978).
11. Rabbitts, T. H., Matthyssens, G. & Hamlyn, P. H. *Nature* **284**, 238–243 (1980).
12. Brenner, S. & Milstein, C. *Nature* **211**, 242–243 (1966).
13. Baltimore, D. *Nature* **248**, 409–411 (1974).
14. Zakut, R., Givol, D. & Mory, Y. Y. *Nucleic Acids Res.* **8**, 453–465 (1980).
15. Zakut, R., Cohen, J. B. & Givol, D. *Nucleic Acids Res.* **8**, 3591–3601, 4840 (1980).
16. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835–851 (1975).
17. Milstein, C., Brownlee, G. G., Harrison, T. M. & Matthews, M. B. *Nature new Biol.* **239**, 117–120 (1972).
18. Tucker, P. W., Marcu, K. B., Slightom, J. L. & Blattner, F. R. *Science* **206**, 1299–1303 (1979).

gene rearrangement additional somatic mutations may occur. Both germ-line substitutions and somatic mutations occur preferentially in the regions coding for CDR.

We thank Mrs Sylvia Wylder for technical assistance. This work was supported in part by a grant from the National Council for Research and Development, Israel, and the DKFZ, Heidelberg, FRG.

19. Matthyssens, G. & Rabbitts, T. H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6561–6565 (1980).
20. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288–294 (1979).
21. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1485–1489 (1978).
22. Auffray, C., Nageotte, R., Chambraud, B. & Rougeon, F. *Nucleic Acids Res.* **8**, 1231–1241 (1980).
23. Kabat, E. A., Wu, T. T. & Bilofsky, H. *Sequences of Immunoglobulin Chains* (NIH publication no. 80-2008).
24. Padlan, E. A., Davies, D. R., Pecht, I., Givol, D. & Wright, C. *Cold Spring Harb. Symp. quant. Biol.* **41**, 627–639 (1977).
25. Padlan, E. A. *Q. Rev. Biophys.* **10**, 35–65 (1977).
26. Capra, J. D. & Kehoe, J. M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 845–848 (1974).
27. Proudfoot, N. J. *Nature* **286**, 840–841 (1980).
28. Bentley, D. L. & Rabbitts, T. H. *Nature* **288**, 730–733 (1980).
29. Proudfoot, N. J. & Maniatis, T. *Cell* **21**, 545–553 (1980).
30. Nishioka, Y., Leder, A. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2806–2809 (1980).
31. Galley, J. A. & Edelman, G. M. *Nature* **227**, 341–348 (1970).
32. Kabat, E. A., Wu, T. T. & Bilofsky, H. *J. biol. Chem.* **252**, 6609–6616 (1977).
33. Smith, H. O. & Birnstiel, M. L. *Nucleic Acids Res.* **3**, 2387–2398 (1976).
34. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
35. Tu, C.-P. D. & Coehn, S. N. *Gene* **10**, 177–183 (1980).

LETTERS TO NATURE

γ Rays from the cosmic ray irradiation of local molecular clouds

M. R. Issa & A. W. Wolfendale

Physics Department, University of Durham,
Durham DH1 3LE, UK

An analysis is given here of the flux of cosmic γ rays of energy above 100 MeV from the directions of 13 local molecular clouds. We show that only in a few clouds does the cosmic ray intensity need to exceed the value near the Earth to explain the γ -ray fluxes. Possible reasons for the few excesses are considered. Our conclusion is important because some of the enigmatic γ -ray 'sources' are identified; furthermore, γ -ray data can be coupled with knowledge about the distribution of gas in the Galaxy to throw light on the problem 'where do cosmic rays come from?'

γ rays are produced when cosmic ray electrons, protons and heavier nuclei collide with atomic nuclei in the interstellar medium (ISM). Insofar as the ISM is not uniformly distributed but contains regions ('clouds') of rather high gas density the γ -ray sky is correspondingly uneven. Of particular interest are those clouds in which the gas, largely molecular hydrogen, is so massive that an excess γ -ray flux can be seen.

Ideally, such clouds could be used to map the distribution of cosmic rays (CR) in the Galaxy, if certain conditions were met: the clouds were of accurately known mass, those clouds selected were 'cosmic ray inert' (that is, generated γ rays only through the interactions of ambient CR and did not contain significant sources of cosmic ray particles) and if the corresponding γ -ray fluxes could be accurately measured. If then clouds which were not 'CR inert' were considered, information could be gained about γ -ray emission from specific sources and also about particular particle acceleration mechanisms.

Clearly, none of these provisos are met to anywhere near the extent necessary for an accurate analysis but an attempt is desirable because of previous conflicting results (for example, for the cloud associated with ρ -Ophiuci; see ref. 1). A serious objection to previous studies was that the γ -ray sources were taken as the starting point; here we start more rigorously with a set of molecular clouds and then search the γ -ray data for

excesses (or deficits) of flux. Hopefully if a large enough sample is taken and the uncertainties in mass and γ -flux are random then the average behaviour will be correct. Unfortunately, there are too few clouds for which data are available to select 'CR inert' clouds. Instead, we must take all clouds for which reasonable mass and distance estimates have been made. The expected γ -ray flux from a cloud of mass M at distance d is estimated from the relation $\Gamma_{\gamma, \text{exp}} = (q/4\pi)(M/d^2)$ where q is the emissivity appropriate to the cosmic ray intensity near the Earth (given by Issa *et al.*¹) and to compare this with the observed γ -ray flux, $\Gamma_{\gamma, \text{obs}}$. Clearly $\Gamma_{\gamma, \text{obs}}/\Gamma_{\gamma, \text{exp}} = I_{\text{cloud}}/I_0$ where I_{cloud} and I_0 are the cosmic ray intensities in the cloud and near the Earth, respectively.

The questions to be asked are: (1) Is the average cosmic ray intensity in the whole sample of local molecular clouds significantly higher than the intensity near the Earth, I_0 ? (Is the average enhancement factor, $F = I_{\text{cloud}}/I_0$, greater than unity?) (2) If there is evidence for $F > 1$ in some clouds and not in others, do they correlate with any known astronomical parameter, such as number of flare-stars, or OB-association?

There is no homogeneous catalogue of masses of molecular clouds determined by the use of one, generally accepted, technique. However, a series of observations by Blitz² give data on seven cloud complexes and these form over half the clouds examined here. The technique was to measure the column density of ^{12}CO and ^{13}CO lines and convert to H_2 density using the technique described by Blitz and Shu³ and elsewhere (which uses the ratio of $N(\text{H}_2)/N(^{13}\text{CO}) = 5 \times 10^5$ of Dickman⁴). Important points are the necessity for an upward correction for the gas associated with each cloud which has too low a density to be recorded directly (usually in the 'halo' around the cloud), and for the effect of channel dilution and non-linearity. Furthermore, the possibility of non-thermodynamic equilibrium in local high density regions is important. Blitz favours upward corrections of between 2 and 3. Here we usually adopt a factor 2 to give our estimated upper limit to the mass (the 'Blitz factor') and take the uncorrected mass as a lower limit. Figure 1 gives the location of the selected clouds and Table 1 the estimated masses and distances.

The local thermodynamic equilibrium (LTE) masses are those given by Blitz. Allowance has been made for excesses of atomic hydrogen and ionized gas as well as molecular hydrogen where data are available. Also given are the approximate areas included in our analysis of the clouds and some astronomical information.

Table 1 Estimated masses and distances of selected clouds

Object	Distance (pc)	LTE mass ($M_{\odot}$)	Mass range ($M_{\odot}$)	Area (deg ²)	Comments	Ref.
Cas OB6	2,000	1×10^5	$1-2 \times 10^5$	4	W3 & W3N, well known H ₂ regions, best known examples of OB star formation. Strong ionization fronts	2
Mon OB1	800	1×10^5	$1-2 \times 10^5$	8	Much recent star formation: T-Tauri	2
Mon OB2	1,600	1×10^5	$3-5 \times 10^5$	4	Rich grouping of early type stars: O stars	2
		$(1 \times 10^5 \text{ for H}_2; 2 \times 10^5 \text{ for H})$				
C Ma OB1	1,100	1×10^5	$1-2 \times 10^5$	4	OB and R associations	2
Orion OB1	450	2×10^5	$2-4 \times 10^5$	20	Many flare stars; T-Tauri common; rich in O-associations	2
Per OB2	350	4×10^4	$4-8 \times 10^4$	24	Some star formation, T-Tauri and O-type	2
Cygnus	1,700	7×10^5	$7-21 \times 10^5$	12	Active star formation—OB associations	2
Gum	400	1.2×10^5	$0.7-3 \times 10^5$	400	Argument about origin of this nebula; considerable ionized gas from supergiants plus stellar winds	28, 29
Cor Aust	150	6.9×10^3	$0.7-1.4 \times 10^4$	8	Contains young stellar objects but efficiency of star formation is very low	15, 30
Taurus	140	1.7×10^4	$1.5-3 \times 10^4$	100	T-Tauri stars distributed throughout the cloud	6, 7
M17	2,300	$>3 \times 10^5$	$3-10 \times 10^5$	4	Cluster of O stars has formed at the edge of the molecular cloud	9
Car Neb	2,700	—	$1.5-3 \times 10^5$ †	5	An unusual object. Star in nebula is one of most luminous in Galaxy. Flared during last century	31
ρ -Oph	160	—	$4.5-9 \times 10^3$ †	16	Cloud very dense; comparatively inert but some star formation	32, 33

† See text.

There is much argument about the origin of the Gum nebula but agreement that there is considerable ionized gas. In view of the proximity of the Vela pulsar and its spillover (in γ rays) we take only the region $b < -6^\circ$. We estimate that $\sim 60\%$ of the gas is in this region (following the H α photograph of Sivan⁵, which suggests that there is more material below the galactic plane than above). The range of masses indicated, which refers to the restricted region, endeavours to bracket the quoted values.

Corona Australis is a useful cloud in that it mirrors ρ -Oph in position and would be expected to have the same cosmic ray intensity in it. Densities in the cores of the sub-clouds are very high. The Blitz factor of 2 possible enhancement in mass has been used but the mass quoted is probably an underestimate.

Taurus 'may be the nearest major aggregate of dust and gas in which star formation is in progress'⁶. It has many T Tauri stars spread throughout the complex. A mass of $1.1 \times 10^4 M_{\odot}$ (within the 0.2 K contour) is quoted by Baud and Wouterloot⁷ for a distance of 113 pc, from observations of OH. For our adopted distance of 140 pc (ref. 6) the corresponding mass is $1.7 \times 10^4 M_{\odot}$ (this is probably an underestimate by $\sim 25\%$ because all the cloud was not surveyed). Another estimate is $2 \times 10^4 M_{\odot}$ using recent ¹²CO data of G. P. Baran (personal communication) for the Taurus and surrounding regions calibrated with the LTE mass obtained by Blitz for the nearby Per OB2 cloud. Bearing in mind the 'Blitz factor' we adopt $1.5-3 \times 10^4 M_{\odot}$ as the likely range for the total mass.

M17 is a bright H II region located on the edge of a large molecular cloud complex which 'may be one of the most massive objects in the Galaxy'⁸. Elmegreen *et al.*⁹ quote a lower limit to the mass of $3.2 \times 10^5 M_{\odot}$, and Elmegreen and Lada⁸ suggest $10^6 M_{\odot}$ for the whole complex. Thus, $3-10 \times 10^5 M_{\odot}$ seems likely.

The star associated with the Carina Nebula (η Carinae) is the most luminous in the Galaxy in the IR ($L \sim 6 \times 10^6 L_{\odot}$). There is strong evidence for an ionization front generated by stars in the Trumpler 14 and 16 clusters, and associated compressed H and H₂ regions¹⁰. The mass is very uncertain because of the lack of CO observations. The estimate of gas in ionized form comes from Dickel's analysis: $1-2 \times 10^5 M_{\odot}$. In view of the inevitable contribution from H and H₂ and the Blitz factor we adopt $1.5-3 \times 10^5 M_{\odot}$ as the likely range.

We previously favoured ρ -Ophiuchi because of its being comparatively inactive: astronomically^{1,11} there are no associated O-stars or bright H II regions¹² although there are several small compact H II regions¹³. There is uncertainty in the mass (see refs 11, 14 for conflicting estimates). Here we consider the following: Myers *et al.*¹² quote $1.8 \times 10^3 M_{\odot}$ (assuming LTE) out to a radial distance of 2.1 pc for H + H₂; Grasdale *et al.*¹³ give $4 \times 10^3 M_{\odot}$ for a somewhat smaller region, Vrba¹⁵ quotes $6.1 \times 10^3 M_{\odot}$ in total. Our previous estimate overall from CO data was $\sim 9 \times 10^3 M_{\odot}$, inclusive of ionized gas (Baart *et al.*¹⁶ find that such gas is common) but our analysis of stellar extinction led to values of about half this. Thus, we adopt $4.5-9 \times 10^3 M_{\odot}$ in the present case.

The γ -ray results will now be considered. The data come from the SAS 2 and COS B experiments. The SAS 2 data comprise the comprehensive tables of Fichtel *et al.*¹⁷ for the two energy ranges 35–100 MeV and $E_{\gamma} > 100$ MeV, and the contours for l 150°–240° and $|b| < 25^\circ$ given by Thompson *et al.*¹⁸. COS B results have not yet been given in tabular form and we have used other sources: the contours for $E_{\gamma} > 70$ MeV of Wills *et al.*¹⁹ which relate to all longitudes and largely to $|b| \geq 15^\circ$, the cut along $b = 0^\circ$ given by the same author and sections for 70–150 MeV, $|b| < 7^\circ$; 150–300 MeV, $|b| < 5^\circ$ and 300–5,000 MeV, $|b| < 4^\circ$ given by Mayer-Hasselwander *et al.*²⁰. The 2CG catalogue of COS B sources of Hermesen²¹ has also been used.

For each cloud we delineate an area surrounding it some 3–4° wider than the cloud (the exact area depending on γ -ray energy and the source of data: tabulated fluxes, contours and flux versus longitude plots for fixed latitude bins) and compare the integrated γ -ray flux in this region with its surroundings, usually by way of adjacent longitude bins for the same latitude range. The adopted flux is the flux in the region less the average of its surroundings. The greater coverage than the 'source' area is to allow for at least some of the γ rays which fall outside the area because of errors of γ -ray direction (the 'point spread function'); a correction is applied for the effect of residual loss, that is, net loss even out of the enlarged area. The excess flux which can be

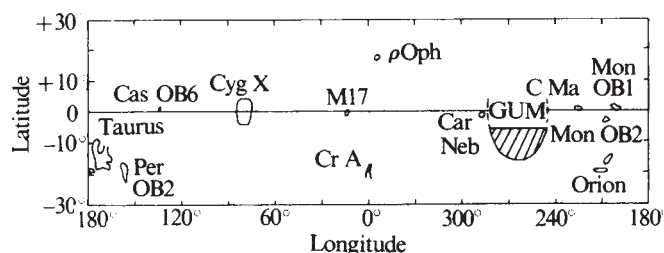
**Fig. 1** Positions and approximate shapes of the molecular cloud complexes.

Table 2 Observed and expected γ -ray fluxes and the enhancement factor

Object	Fluxes ($10^{-6} \text{ cm}^{-2} \text{ s}^{-1}$) ($E_\gamma > 100 \text{ MeV}$)				F needed to give measured γ -flux
	SAS 2	COS B	Best flux estimate	Expected flux	
Cas OB6	0.5	1.0 (if 2CG135+01)	1.0	0.07–0.14	7–14
Mon OB1	<0.3	0.5	0.3	0.4–0.9	0.3–0.8
Mon OB2	<0.2	<0.5	<0.4	0.3–0.5	<1.3
CMA OB1	<0.5	0.5	0.5	0.23–0.5	1–2
Orion OB1	1.0	1.4 (ref. 22)	1.4	2.8–5.6	0.25–0.5
Per OB2	<0.8	—	<0.8	0.9–1.8	<0.9
Cygnus	4.0	3.8 (if 2CG075–00 and 2CG078+01)	3.8	0.7–2.1 2.7*	1.8–5.4
Gum	4.2	—	4.2	1.2–5.3	0.8–3.5
Cor Aust	0.4	<0.4	0.4	0.9–1.8	0.2–0.4
Taurus	0.5	—	0.5	2.1–4.2	0.12–0.24
M17	<0.6	1.0 (if 2CG13+00)	1.0	0.16–0.53	1.9–6.3
Car Neb	<2.2	1.6 (if 2CG288–00)	1.6	0.06–0.12	13–27
ρ -Oph	<0.7	1.1 (if 2CG353+16)	1.1	0.5–1.0	1.1–2.2

* This value is from the analysis of Protheroe *et al.*³⁴ using column densities but converting to $q/4\pi = 2.2 \times 10^{-26} \text{ s}^{-1} \text{ sr}^{-1}$ as adopted here.

attributed to the cloud is estimated using the above sources of data. Where the flux is negative, upper limits at the 1σ level are derived.

This cannot be an accurate procedure as there are many problems associated with inhomogeneous backgrounds and poor statistics; however, it seems to be objective and we can see no reason for the eventual average (or better, the median) value of the flux enhancement factor, F , being either underestimated or overestimated.

To improve accuracy, the data are combined to give an estimate of the flux above one energy level: 100 MeV. To this end, the SAS 2 fluxes for $35:100 \text{ MeV}$ are multiplied by 0.68 to give the equivalent flux above 100 MeV.

Table 2 gives the observed and our estimated γ -ray fluxes and the enhancement factor, F . The 2CG notation refers to the COS B source catalogue of Hermsen²¹ and these fluxes are taken where the 'sources' are in such positions as to be not inconsistent with the clouds. Positive identification is not possible, however, and there are doubts in some cases. The COS B flux value for Orion comes from Caraveo *et al.*²² which relates to $E_\gamma > 70 \text{ MeV}$ ($\approx 2 \times 10^{-6} \text{ cm}^{-2} \text{ s}^{-1}$); conversion by the spectral shape factor yields $\approx 1.4 \times 10^{-6} \text{ cm}^{-2} \text{ s}^{-1}$ above 100 MeV. This value is close to the $1.0 \times 10^{-6} \text{ cm}^{-2} \text{ s}^{-1}$ derived from analysis of SAS 2 data¹¹. (Note that Caraveo *et al.* derive $F \approx 1$ by adopting a mass of $1.2 \times 10^5 M_\odot$, derived from Galaxy count studies.) No value is given for the flux for those regions where no data have been reported (or at least there are no data over a wide enough $\Delta l \times \Delta b$ range to allow flux estimates to be made).

Figure 2 shows the values of F listed in Table 2 plotted against the distance from the Earth, d . Ideally F should be plotted against distance from the axis of the nearest spiral arm (in an effort to search for an arm-interarm contrast in cosmic ray intensity) but the positions of the axes are not sufficiently well known. Although not necessarily expecting a dependence of F on d over the range indicated we do expect a loss of reliability at large d because of the possibility of confusion from other clouds and this will be particularly severe when looking roughly along the axis of a spiral arm (specifically for Car Neb and, to a lesser extent, Cygnus). In fact, the 2CG COS B source catalogue lists 2CG288–00 (the Car Neb source) and 2CG284–00 nearby as 'could be an extended feature'.

Taking the data in Fig. 2, the average value of F is ~ 1.3 , weighting according to the errors indicated. Although this estimate is not very precise in view of there being errors in γ -ray

flux, which have not been included, the extra errors should be symmetrical and the estimate should not be too far out. It seems likely, then, that most of the clouds do not have much enhancement ($F \leq 2$) of cosmic ray intensity within them. In this sense the results are consistent with a previous analysis²³ in

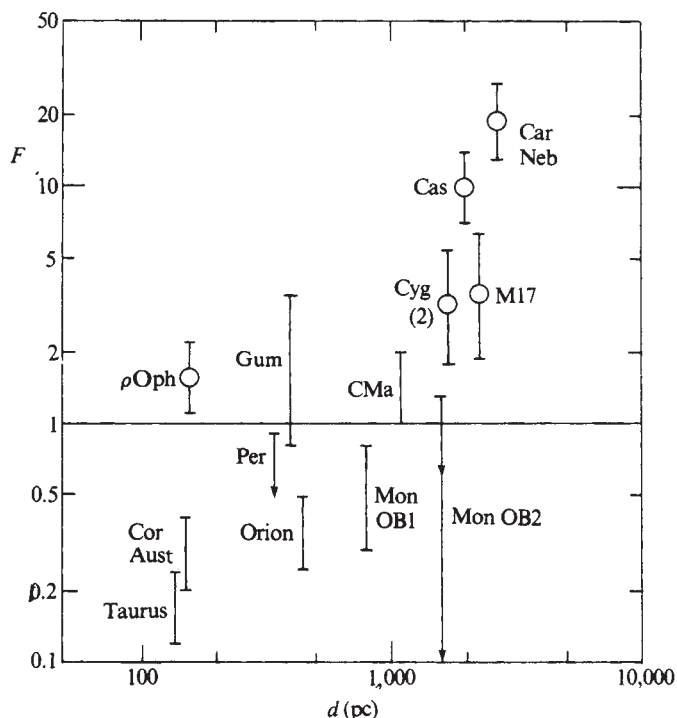


Fig. 2 F -values for the molecular clouds indicated. The ranges shown come from uncertainty in cloud mass; no uncertainty in γ -ray flux has been allowed for. The flux errors are expected to be roughly symmetrical so that the best-estimate F -values will not change. The abscissa is the distance from the Earth. The circles on some of the error bars indicate clouds which may be identified with γ -ray sources in the 2CG catalogue; the 2CG γ -ray fluxes have been adopted in the calculations (the (2) against Cygnus denotes two COS B sources). The astronomical data of Table 1 are probably best for Orion OB1, Per OB2 and Mon OB2 (J. Meaburn, personal communication).

which it is claimed that over half of the 2CG sources could be explained as being due to CRI molecular clouds, the analysis being made on a statistical basis.

Despite the large uncertainties there are some features which are worthy of further attention.

(1) The cloud complexes within 1 kpc of the Earth mainly have $F < 1$. Some of these have many T-Tauri stars and this indicates that the stellar winds from T-Tauri stars are not the seat of much cosmic ray acceleration or, at least, if they do accelerate particles then they escape rapidly from the environment of the stars (the likelihood of CR acceleration by stellar winds was first pointed out by Cassé and Paul²⁴). It is possible, but not certain, that some of the low values of F relate to the position of the clouds—they may be preferentially in interarm regions. Beyond 1 kpc (where T-Tauri stars cannot be identified) F -values are generally higher and there seem to be frequent OB-associations. The significance of this is not clear because of selection phenomena; the relevance of OB-associations to γ -ray sources has already been discussed by Montmerle²⁵ who required nearby supernovae ('SNOB'S') to augment I_{cloud} . We do not feel much increase in flux is needed.

(2) The only clouds which seem to need F -values significantly greater than unity are Car Neb and Cas OB6, the differences from unity for the others can be attributed in part to errors of γ -ray flux, and of cloud mass and to the likelihood of small variations of F from place to place on a large scale.

The excess in Car Neb may be genuine, with particle acceleration occurring in a manner such as described by Montmerle²⁶. However, several features need stressing. The first concerns the fact that the direction of Car Neb is congested in the sense that it is along a spiral arm and other clouds may also be contributing within the angular resolution of the instrument. The mass of cloud needed (at the distance of Car Neb) is $\approx 3 \times 10^6 M_{\odot}$; although large, this is not an impossibly large mass for a single giant molecular cloud. The inclusion of other clouds would relax the mass requirement considerably. Also the identification of 2CG288-00 with Car Neb may be erroneous. The source could be, for example, a pulsar in this direction; pulsars do emit γ rays.

There is the same possibility of mis-identification with Cas OB6. Coincidence of more than one cloud is unlikely, however, because the direction is not along an arm. What is interesting is the fact that Cas OB6 is in a well developed arm—the Perseus arm—and the possibility of some general CR intensity enhancement arises.

Although all the data—both γ -ray and astronomical—are poor such evidence as there is does not call for much CR enhancement in most of the molecular clouds within 1 or 2 kpc of the Earth ($\langle F \rangle \approx 1.3$ for all the clouds in Fig. 2 and even nearer unity for $d < 2$ kpc). Unidentified molecular clouds irradiated by CR and situated beyond 2 kpc are likely contenders for the so-called γ -ray sources and, indeed, there are indications²³ that $\sim 60\%$ of the COS B sources can be explained in this way.

Thus the studies of the large scale distribution of γ -ray flux, in which unresolved discrete sources have been neglected and it has been concluded that there are large scale cosmic ray gradients in the Galaxy, should be valid (see, for example, ref. 27).

We do not suggest that the sources of some of the low energy cosmic rays are not in molecular clouds. They could well be, but diffusion will probably distribute the particles so rapidly that those interacting in a particular cloud are mainly from the general ambient cosmic radiation.

We thank Dr L. Blitz for the copy of his thesis and Dr A. W. Strong for helpful comments. Dr J. Meaburn and Dr S. M. Scarrott suggested improvements to the manuscript.

Received 18 February; accepted 2 June 1981.

- Issa, M. R., Strong, A. W. & Wolfendale, A. W. *J. Phys. G: Nucl. Phys.* **7**, 565 (1981).
- Blitz, L. in *Giant Molecular Clouds in the Galaxy* (eds Solomon, P. M. & Edmunds, M. G.) 1 (Pergamon, Oxford, 1980).
- Blitz, L. & Shu, F. H. *Astrophys. J.* **238**, 148 (1980).
- Dickman, R. L. *Astrophys. J. Suppl.* **37**, 407 (1978).

- Sivan, J. P. *Astr. Astrophys. Suppl.* **16**, 163 (1974).
- Elias, J. H. *Astrophys. J.* **224**, 857 (1978).
- Baud, B. & Wouterloot, J. G. *Astr. Astrophys.* **90**, 297 (1980).
- Elmegreen, B. G. & Lada, C. J. *Astr. J.* **81**, 1089 (1976).
- Elmegreen, B. G., Lada, C. J. & Dickinson, D. F. *Astrophys. J.* **230**, 415 (1979).
- Dickel, H. R. *Astr. Astrophys.* **31**, 11 (1974).
- Wolfendale, A. W. *IUPAP/IAU Symp.* No. 94, 309.
- Myers, P. C., Ho, P. T. P. & Schneps, M. H. *Astrophys. J.* **220**, 864 (1978).
- Grasdalen, G. L., Strom, K. M. & Strom, S. E. *Astrophys. J. Lett.* **184**, L53 (1973).
- Paul, J. A., Cassé, M. & Montmerle, T. *IUPAP/IAU Symp.* No. 94, 325.
- Vrba, F. J. *Astr. J.* **82**, 198 (1977).
- Baart, E. E., de Jager, G. & Mountfort, P. I. (in the press).
- Fichtel, C. E. *et al. NASA Tech. Mem.* 79656 (GSFC, Greenbelt, 1978).
- Thompson, D. J., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Lamb, R. C. *NASA Tech. Mem.* X-662-76-198 (GSFC, Greenbelt, 1976).
- Wills, R. D. *et al. Non-solar Gamma Rays* (COSPAR) (eds Cowdik, R. & Wills, R. D.) 43 (Pergamon, Oxford, 1980).
- Mayer-Hasselwander, H. A. *et al. Ann. N.Y. Acad. Sci.* **336**, 211 (1980).
- Hermesen, W. thesis, Univ. Leiden (1980).
- Caraveo, P. A. *et al. Astr. Astrophys.* **91**, L3 (1980).
- Li, T. P. & Wolfendale, A. W. *Astr. Astrophys.* (in the press).
- Cassé, M. & Paul, J. A. *Astrophys. J.* **237**, 236 (1980).
- Montmerle, T. *Astrophys. J.* **231**, 95 (1979).
- Montmerle, T. *Phil. Trans. R. Soc.* (in the press).
- Issa, M. R., Riley, P. L., Strong, A. W. & Wolfendale, A. W. *Nature* **287**, 810 (1980).
- Reynolds, R. J. *Astrophys. J.* **203**, 151 (1976).
- Brandt, J. C., Stecher, T. P., Crawford, D. L. & Maran, S. P. *Astrophys. J. Lett.* **163**, L99 (1971).
- Rossano, G. S. *Astr. J.* **83**, 234 (1978).
- Turner, D. G. & Moffat, A. F. J. *Mon. Not. R. astr. Soc.* **192**, 283 (1980).
- Bertaud, F. C. *Astrophys. J.* **128**, 533 (1958).
- Whittem, D. C. B. *Mon. Not. R. astr. Soc.* **168**, 371 (1974).
- Protheroe, R. J., Strong, A. W. & Wolfendale, A. W. *Mon. Not. R. astr. Soc.* **188**, 863 (1979).

Collisional amplification of density fluctuations in Saturn's rings

J. Lukkari

Department of Astronomy, University of Oulu, Oulu, Finland

The encounter of Voyager 1 with Saturn in November 1980 revealed the ringlet structure of its rings. As the theoretical examination of the collisional evolution of keplerian systems^{1,2} had predicted such a structure in dense matter as a consequence of amplified fluctuations in density, some new computer simulations have now been carried out to check this effect. These simulations actually led to a strong, irreversible growth in the density maximum.

Let a keplerian system consist of frictionless, spherical and identical particles revolving around the central body along elliptic orbits and having the radius σ . Let a denote the semi-major axis, e the eccentricity, i the inclination, γ Newton's constant and m the mass of the central body. The local value of optical thickness is τ .

Solid materials have a restitution coefficient, α , which decreases with the velocity of impact, v . If α_0 denotes the particular value of α which corresponds to the root-mean-square velocity of impacts, the system adjusts itself to a state in which²

$$\frac{9\pi^2(19\alpha_0 - 13)}{64(1 - \alpha_0^2)(3 - \alpha_0)^3} = g^2\tau^2, \quad \alpha_0 = \alpha(\sqrt{v^2}) \quad (1)$$

The weight function g is a correction for the finite volume of the particles; if the space density is low, we have $g = 1$. As the space density depends on the geometric thickness of the system, its optical thickness is not restricted by the value of g . If equation (1) is valid, the radial component of the two-dimensional flux vector which represents the transport of particles on the equatorial plane of the system can be calculated from²

$$F^r = -\frac{8}{9\pi^2\sigma^2\sqrt{\gamma ma}} \frac{\partial}{\partial a} [g\tau^2 a^2 T(1 - \alpha)\sqrt{(1 + \alpha)(3 - \alpha)}] \quad (2)$$

The temperature-like quantity T is approximately equal to $3v^2/4$. According to equation (1), $\sqrt{v^2}$ is a decreasing function of

the optical thickness, as has also been found by Goldreich and Tremaine³. If τ is high enough, this effect produces a decrease in the function inside the square brackets in equation (2). For example, if $\alpha = (1 + v/v_0)^{-1}$ ($v_0 = \text{constant}$) and $g = 1$, this function has its largest value at $\tau = 0.75$. If τ exceeds it, the flux is directed to the density maximum. If τ is small, the same function increases with τ and the flux is directed away from the maximum density. In particular, as $\tau \rightarrow 0$ at the boundaries of the system, these expand with a macroscopic velocity $\pi\sigma^2 F^2/\tau \sim -\partial\tau/\partial a$. As a whole, the system therefore expands. This agrees with Brahic's⁴ conclusion. As Brahic derives the expansion from the conservation of the total angular momentum of the whole system and from a simultaneous decrease in its total energy, his result cannot be generalized for parts of a system. Their angular momentum may increase or decrease according to the interaction with adjacent regions. A local contraction of matter is therefore possible.

In the present report we are not studying the general expansion, but are testing the reality of the amplification of fluctuations which is predicted by equation (2) at high values of τ . Because Saturn's rings exist, the time scale of the expansion must be long or some mechanism must prevent it. The formation of ringlets has a shorter time scale. Once the fluctuations have begun to grow, the gradient in equation (2) becomes progressively steeper and the flux increases until the finite size of particles stops the process.

Previous simulations⁵ with a constant coefficient of restitution indicate that the validity of theoretical predictions with $g = 1$ is restricted by the condition $\sqrt{i^2} \gg 4\sigma/a$. On the other hand, in a ring-shaped system of radial width w , the gradients in density and in the distribution functions are too high unless $w/2 \gg a\sqrt{\epsilon^2}$, $a\sqrt{i^2}$. This follows from the limits of validity for theoretical expressions^{2,5}. As $\sqrt{i^2/\epsilon^2} \approx 2/3$ in the equilibrium state⁵, we can also write both conditions in the form $w \gg 2a\sqrt{\epsilon^2} \gg 12\sigma$. If the sign $\gg$ is interpreted to mean the ratio 5:1, the above condition can be written $w \geq 300\sigma$. The number of particles in a ring of optical thickness τ is $2aw\tau/\sigma^2$. If $\tau > 0.75$ as in the above-mentioned numerical example, and if $\sigma/a = 0.001$, a system of 450,000 particles is needed to satisfy the condition $w \geq 300\sigma$. The large number of particles prevents the direct simulation of dense matter (Saturn's rings). Because the dependence between τ and α is of a statistical nature, one cannot even use a Monte Carlo method, which only increases the number of particles but does not change their statistics. We must therefore use an indirect method.

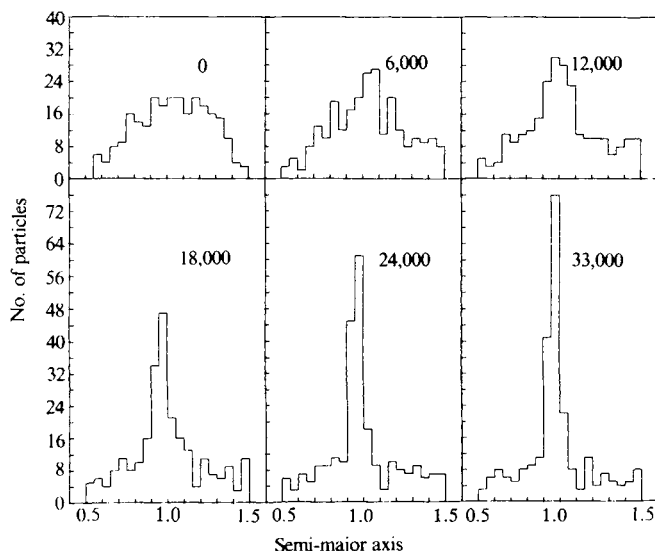


Fig. 1 Radial distribution of particles during one simulation. The number of collisions is given beside each histogram. The corresponding times are 0, 132,000, 247,000, 335,000, 403,000 and 484,000 in units in which $\sigma = 0.001$ and $\gamma m = 1$.

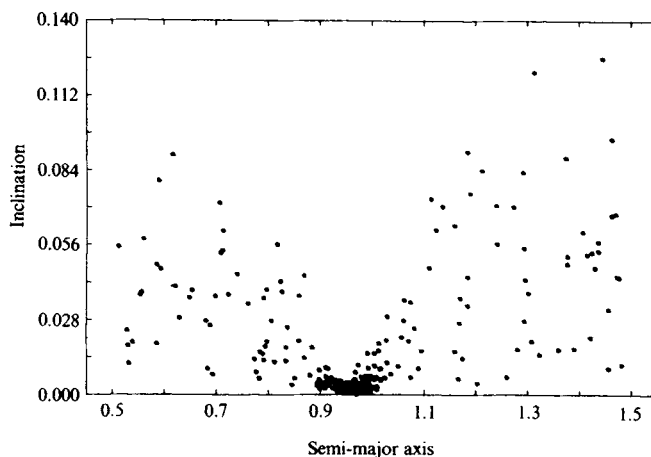


Fig. 2 Inclinations of the orbits as a function of semi-major axis after 30,000 collisions.

In rarefied matter, equation (1) gives the result $\alpha_0 = 13/19$, which has also been confirmed by a computer simulation⁵. Applying this result we can artificially force a rarefied system to behave in the same manner as a dense one. For this purpose it is sufficient to use a coefficient of restitution which depends on both v and τ and produces the result $d\sqrt{v^2}/d\tau < 0$, when the system adjusts itself to $\alpha_0 = 13/19$. This corresponds to equation (1). The function $\alpha(\tau, v)$ can be so chosen that the simulation produces the same flux as equation (2) (but not with the same value of v^2) if the elastic properties of matter are specified. Because we lack this information we can choose $\alpha(\tau, v)$ freely.

In our simulations the restitution coefficient was

$$\alpha = \frac{1}{1 + k(\tau^3 + \tau_0^3)v} \quad k, \tau_0 = \text{constant} \quad (3)$$

In units $\gamma m = 1$, the constants of equation (3) are $k = 1.91 \times 10^{12}$ and $\tau_0 = 7.52 \times 10^{-5}$. The number of particles in zones of width 0.1 was used to calculate the local values of τ . The number of particles is 250 and their radius is 0.001. The initial distribution of τ is proportional to $(1.5 - a)(a - 0.5)$ while the initial values of $\sqrt{\epsilon^2}$ and $\sqrt{i^2}$ are 0.049 and 0.032, respectively. As the number of particles is restricted, we cannot simulate a sequence of adjacent density fluctuations. It is therefore necessary to pick up just one condensation and to follow its evolution. To simulate the pressure of adjacent regions, the particles which were outside the initial zone $0.5 < a < 1.5$ after a collision were brought back into it. The percentage of such impacts was usually $< 1\%$. A slow expansion of this zone would not change the result. The same process can also be interpreted as replacing the unknown mechanism which prevents the general expansion of boundaries in saturnian rings. If the post-collisional semi-major axis was < 0.5 , it was replaced by $1 - a$. If it was > 1.5 , it was replaced by $3 - a$. This has no primary effect on the development of the central spike, because the maximum change of semi-major axis in a collision is usually not larger than $2\epsilon a$. Figure 1 shows the evolution in the radial distribution of the particles in one of the simulations. The increase in density is clearly seen after 6,000 collisions and it continued when the simulation was stopped after 33,000 collisions. Contraction continues although the inclination was at its densest part $\sim 2\sigma/a$, which has previously led to an increasing dispersion of the system⁵. The system thus achieved a state in which α was not able to reach its quasi-equilibrium value $13/19$ but remained smaller.

Figure 2 represents the distribution of inclinations as a function of the semi-major axis after 30,000 collisions. As can be seen the large inclinations in rarefied matter between the ringlets may well cause the observed thickness of Saturn's rings, whereas the thickness of the ringlets is much smaller.

Extremely rarefied matter outside Saturn's rings or self-gravitation or some external perturbation is needed to prevent the extreme borders of the whole sequence of ringlets from dispersing into space. If this is assumed, then the simulations show in accordance with the theory^{1,2} that an initially almost homogeneous ring of dense matter is spontaneously divided into ringlets.

Received 2 March; accepted 17 June 1981.

1. Hämeen-Anttila, K. A. *Astrophys. Space Sci.* **56**, 421–430 (1978).
2. Hämeen-Anttila, K. A. *Astrophys. Space Sci.* **58**, 477–519 (1978).
3. Goldreich, P. & Tremaine, S. *Icarus* **34**, 227–239 (1978).
4. Brähic, A. *Astr. Astrophys.* **54**, 895–907 (1977).
5. Hämeen-Anttila, K. A. & Lukkari, J. *Astrophys. Space Sci.* **71**, 475–497 (1980).

The ice layer in Uranus and Neptune—diamonds in the sky?

Marvin Ross

University of California, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

Many of the current models of Uranus and Neptune postulate a three-layer structure, consisting of an inner rocky core, a middle 'ice' layer of fluid, H_2O , CH_4 , NH_3 and an outer hydrogen-helium layer of solar composition¹. The estimated pressures and temperatures of the ice layer ranges from about 6 Mbar and 7,000 K at the inner core-ice boundary, to ~0.2 Mbar and 2,200 K at the outer ice/hydrogen-helium boundary. I point out here that shockwave experiments on these liquids^{2–5}, as well as theoretical studies^{6–7}, imply that the H_2O and NH_3 in the ice layer are almost totally ionized and the CH_4 has been pyrolysed to carbon, possibly in the metallic or diamond form^{8,9}.

In recent years shock-wave experiments have been carried out on all of the fluid constituents of the outer planets except helium. In a shock-wave experiment, the change in pressure, density, and energy, on compression are measured, but the temperature must be computed from the equation of state. A series of experiments starting from the same initial conditions (that is, the normal liquid) and spanning a range of final pressures and densities is referred to as a Hugoniot curve. Figure 1, in comparing some of these data for the 'ices' with an isentrope calculated for Uranus (Neptune is very similar) demonstrates the significance of the data to theoretical modelling studies of these planets. For methane the experimental data, only some of which is shown, consists of a principal Hugoniot up to 0.45 Mbar, plus one data point reflected from 0.23 to

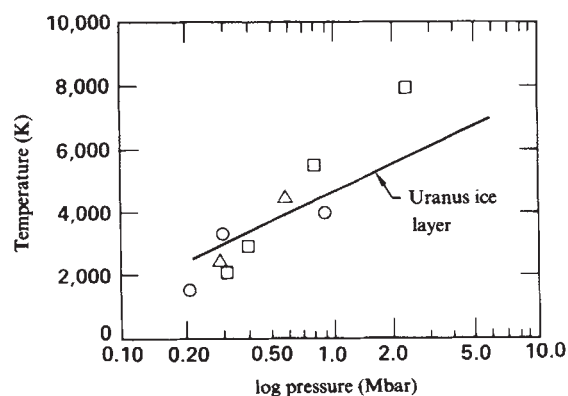


Fig. 1 Plot of some of the experimental shock-wave data for water ($\square$), ammonia ($\triangle$) and methane ($\circ$) in the range of temperatures and pressures believed to exist in the Uranus (and Neptune) ice layer.

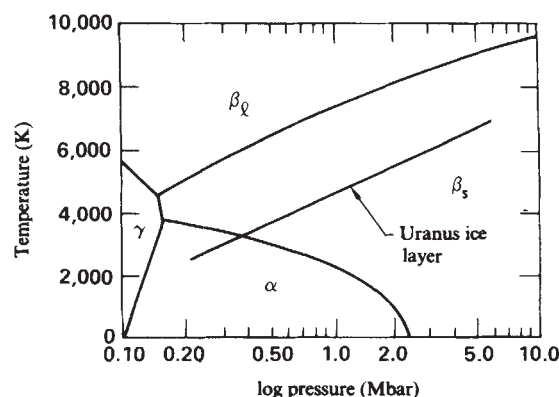


Fig. 2 The Uranus (and Neptune) ice layer (straight line) compared with pressure and temperatures for the diamond phase diagram of Grover⁸. α , Diamond; γ , graphite; β_S , metallic solid; β_L , metallic liquid. Van Vechten's⁹ phase diagram predicts that in the region of the isentrope, above ~0.4 Mbar, carbon will be a liquid metal.

0.91 Mbar. It has recently been shown that an apparent deviation from predicted molecular behaviour could be explained by the tendency for CH_4 to dissociate in these conditions⁶. These conclusions were based on chemical equilibria calculations which predicted that above 0.20 Mbar and 2,000 K, methane is converted into elemental carbon and molecular hydrogen. Recent theoretical calculations on Hugoniot of many hydrocarbons indicate that shock heating induces breaking of the C–H bonds, and the compression encourages condensation of the dissociated carbon atoms into a residue⁷. If we assume that the hydrocarbons have been completely converted into a mixture of condensed carbon and molecular hydrogen, we can use known equations of state for each of these materials to compute high-density Hugoniot that are in excellent agreement with the experimental hydrocarbon data. The results of similar calculations for molecular methane⁶ also agree with the experimental data suggesting that the final product is also hydrogen and a carbon residue.

The Uranus isentrope is shown again in Fig. 2, this time plotted with the recently proposed carbon phase diagrams of Grover⁸. Van Vechten⁹ had previously proposed a carbon phase diagram. Although they differ somewhat in detail above 3,000 K and 0.4 Mbar, they do agree in predicting that below these conditions carbon will exist in the well-known 4-coordinated diamond form. Above these conditions Grover predicts that carbon along the planetary isentrope will be in a metallic solid phase whereas Van Vechten predicts a liquid metallic phase. In Van Vechten's theory the close-packed metal becomes liquid at temperatures well below those of the planetary ice layer. Currently, hydrogen is believed to become a metal somewhere between 2.0 and 5.0 Mbar. Consequently, over much of the higher pressure range of the isentrope the hydrogen freed from the molecular methane will also be metallic and may form a metallic carbon hydrogen mixture.

As indicated in Fig. 1, shock-wave data on water and ammonia, including electrical conductivities, have been measured over part of the Uranus (Neptune) pressure-temperature range. The electrical conductivity data shows that above 0.2 Mbar and about 2,000 K the conductivities become constant at ~20–30 S cm^{-1} , as if the processes leading to ionization have become saturated. Water has apparently become fully ionized³. For ammonia similar results have been observed³. Consequently we must conclude that in the Uranus and Neptune ice layer H_2O and NH_3 are not molecular but ionic, and that the carbon in methane has been converted to a diamond or metallic phase with molecular or metallic hydrogen. Shock-wave data on diamond and graphite shows that the carbon condensate at a few megabars will be denser than NH_3 or H_2O and, unless it is highly soluble in the fluids, may separate out and sink below to form a denser layer.

The electrical properties of these ices are of great interest because they can lead to an explanation or a prediction of a magnetic field. The motion of charged particles trapped in such magnetic fields generates radio waves. There is one report of possible radio signals from Uranus, but there have been no similar reports from Neptune¹⁰. The significance of a conducting layer of carbon may be very great because it is estimated that the mass of available carbon in Uranus and Neptune is 17% of the total planetary mass¹.

Work performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

Received 4 May; accepted 2 June 1981.

- Hubbard, W. B. & McFarland, J. J. *J. geophys. Res.* **85**, 225 (1980).
- Mitchell, A. C. & Nellis, W. J. in *High Pressure Science and Technology* Vol. 1 (eds Timmerhans, K. D. & Barber, M. S.) 428 (Plenum, New York, 1980).
- Mitchell, A. C., Kovel, M. I., Nellis, W. J. & Keeler, R. N. in *High Pressure Science and Technology* Vol. 2 (eds Vodar, B. & Marteau, P.) 1048 (Pergamon, Oxford, 1980).
- Nellis, W. J., Mitchell, A. C., Ross, M. & van Thiel, M. in *High Pressure Science and Technology* Vol. 2 (eds Vodar, B. & Marteau, P.) 1043 (Pergamon, Oxford, 1980).
- Nellis, W. J., Ree, F. H., van Thiel, M. & Mitchell, A. C. *J. chem. Phys.* (submitted).
- Ross, M. & Ree, F. H. *J. chem. Phys.* **73**, 6146 (1980).
- Ree, F. H. *J. chem. Phys.* **70**, 974 (1979).
- Grover, R. J. *J. chem. Phys.* **71**, 3824 (1979).
- Van Vechten, J. A. *Phys. Rev.* **B7**, 1479 (1973).
- Brown, W. L. *Astrophys. J. Lett.* **207**, L209 (1976).

A refined estimate of the fine structure constant

John Ellis & D. V. Nanopoulos

CERN, CH 1211 Geneva 23, Switzerland

Quantum electrodynamics (QED) is an immensely powerful theory whose validity extends over distance scales from 10^{-16} cm to light years. But it is unsatisfactory for many reasons, not least of which is the arbitrariness of the electromagnetic coupling strength, the fine structure constant α , and the incomprehensible commensurability of the electric charges of different particles. When extending QED to distances smaller than 10^{-16} cm it is common to combine it first with the weak interactions¹⁻³ and subsequently with the strong interactions into a grand unified theory (GUT) of all elementary particle interactions^{4,5}. The incorporation of QED and the other fundamental interactions into a GUT is only possible if these interactions and the particles experiencing them satisfy some non-trivial constraints⁶. It is well-known that the incorporation of QED into a simple unifying group imposes charge quantization^{4,5}. Stability of baryons in GUTs requires the grand unification energy scale m_X to be $\geq 10^{14}$ GeV. Couplings in renormalizable field theories such as QED evolve logarithmically with energy, and for the effective α to be less than unity—necessary if the GUT is to make sense—then $>1/25$ (ref. 7). We point out here that within the same standard set of grand unifying assumptions used previously, a much more refined estimate of the fine structure constant is possible: $1/120 \geq \alpha \geq 1/170$. We use only the general observations⁶ that light fermions can be grouped into 'families' or 'generations' of 15 helicity states, that the strong interactions become strong of a scale of the order of 1 GeV, and that any grand unification energy scale must lie between 10^{14} (for baryons to be stable enough) and 10^{19} GeV (the energy at which gravity becomes of unit strength). We know of no other theoretical framework⁸ for particle interactions which can provide such convincing and restrictive constraints on α , and we find it hard to believe that the experimental value of $\alpha = 1/137$ is a coincidence.

We now recall some basic features of the GUTs before deriving our upper and lower limits of α (for more complete descriptions see refs 6, 7, 9). The strong interactions are described by a non-Abelian SU(3) gauge group, while the weak and electromagnetic interactions are described by a theory¹

based on the group SU(2) $\times$ U(1). The ratio of the gauge coupling constants g_2 and g_1 in this theory is arbitrary, and is parametrized by $\sin^2 \theta_w$, and the fine structure constant $\alpha = \alpha_2 \sin^2 \theta_w$ where $\alpha_2 \equiv g_2^2/4\pi$. The strong, weak and electromagnetic interactions act on fundamental fermions, the quarks and leptons, which seem to be arranged^{4,5} in sets of 15 helicity states called 'families' or 'generations'. GUTs combine the different SU(3), SU(2) and U(1) gauge interactions into a simple group G, and assign the fermions to 15 dimensional representations of G. It has been emphasized⁶ that it is not trivial that the observed fermions fall into representations of 'low energy' symmetries which can be grand unified.

Grand unification introduces new gauge interactions between the fermions which violate baryon and lepton number conservation. They are mediated by heavy gauge bosons of masses m_X heavier than 10^{14} GeV if protons and bound nucleons are not to have lifetimes shorter than the present experimental lower limit¹⁰ of 10^{30} yr. In their present formulations, GUTs do not include gravitation^{11,12}. As quantum gravity effects (such as one-graviton exchange) become important at energies comparable with the Planck mass of 1.2×10^{19} GeV, it follows that the grand unification mass scale must be $< 10^{19}$ GeV if the GUT concept is to make sense.

The grand unification mass scale can be estimated¹³ using the renormalization group^{14,15}, the idea being that while the SU(3), SU(2) and U(1) coupling constants are equal above this scale, below it they have different logarithmic dependences on the effective energy scale at which they are measured. This is how¹³ the SU(3) interactions can become strong at energies of the order of 1 GeV, while the SU(2) $\times$ U(1) weak interactions can be weak at low energies. For the GUT philosophy to be viable, the 'low-energy' weak interactions must be neither too strong—in which case grand unification would occur at too low a mass scale and baryons would decay too fast—nor too weak—in which case grand unification would not occur before the Planck mass and it would have been wrong not to have included gravity.

The fine structure constant $\alpha = \alpha_2 \sin^2 \theta_w$, where $\sin^2 \theta_w$ is determined by GUT to be $3/8$ in the symmetry limit above the grand unification scale^{4,5,13}. (The variation of $\sin^2 \theta_w$ below this scale has been computed^{13,16,17}, and the value of 0.21–0.22 found at low energies lies within 1 s.d. of the present experimental value (I. Liede and M. Roos, personal communication).) Knowing the symmetry value of $\sin^2 \theta_w$ we can translate the 'grand unifiability' restriction on the weakness of the weak interactions into upper and lower limits on α itself. In the leading order approximation, the renormalization group equations at energy scales Q between 100 GeV ($\approx m_{W^\pm}, m_{Z^0}$) and the grand unification scale m_X tell us that

$$\frac{1}{\alpha(Q)} \approx \frac{11}{\pi} \ln(m_X/Q) + \frac{8}{3} \frac{1}{\alpha_3(Q)} \quad (1)$$

if we only include GUT representations of fermions and gauge bosons. From renormalization, if $\alpha_3(Q)$ is of the order of 1 on a typical hadronic energy scale then it should be of the order of 0.1–0.2 when $Q \sim 100$ GeV (refs 18–21). There is also a change in the effective $\alpha(Q)$ between its value at the Thompson limit $Q = 0$, which is where we define and measure α , and its value at $Q \approx 100$ GeV:

$$\frac{1}{\alpha} \approx \frac{1}{\alpha(100 \text{ GeV})} + 9 \quad (2)$$

If we use these low energy renormalizations of $\alpha_3(Q)$ and $\alpha(Q)$ and require $10^{14} \text{ GeV} < m_X < 10^{19} \text{ GeV}$ in equation (1) we find $1/119 < \alpha < 1/174$ which we have rounded off to the values quoted earlier. Including a possible light Higgs boson into the leading order renormalization group equations changes these bounds by $\sim 1\%$.

Any theorist working on grand unification would have known that

$$\frac{m_X}{\Lambda_{\text{QCD}}} = \exp \left\{ \frac{0(1)}{\alpha} + 0(1) \times \ln \alpha + 0(1) \times \alpha^0 + \dots \right\} \quad (3)$$

and put in the known value of α to determine m_X in minimal GUTs within a factor of 10 or less²². However, to our knowledge no-one has inverted the procedure and pointed out how tightly constrained a grand unifiable α must be. Indeed the previous⁷ upper limit on α from GUTs was much weaker than the present one. The most precise estimates²² of m_X have gone beyond our leading order formula (1) to include all the $\ln c$ and 0 terms of equation (3). As we intend to convey a qualitative message, we will not repeat all those calculations here. However, we emphasize that they should enable one to make an even more refined estimate of α if baryon decay is ever observed. This is because the baryon lifetime²² is very sensitive to the value of m_X :

$$\tau_{\text{baryon}} \approx 8 \times 10^{30 \pm 1} \text{ yr} \times \left(\frac{m_X}{6 \times 10^{14} \text{ GeV}} \right)^4 \quad (4)$$

Hence a determination of τ_{baryon} should enable us to fix m_X within a factor of 1.5 or so. This 10-fold improvement in bounding m_X would in turn enable us to estimate α^{-1} an order of magnitude more precisely say to ± 3 . The fact that α lies closer to the upper limit we have established gives us hope that m_X is closer to 10^{14} GeV than to 10^{19} GeV, and more detailed estimates in popular GUTs strongly suggest²² that the baryon lifetime is sufficiently short to be detectable in the forthcoming generation of experiments.

Received 6 May; accepted 5 June 1981.

1. Glashow, S. L. *Nucl. Phys.* **22**, 579 (1961).
2. Weinberg, S. *Phys. Rev. Lett.* **19**, 1264 (1967).
3. Salam, A. *Proc. 8th Nobel Symp.* Stockholm (ed. Svartholm N.) 367 (Almqvist and Wiksells, Stockholm, 1968).
4. Georgi, H. & Glashow, S. L. *Phys. Rev. Lett.* **32**, 236 (1974).
5. Pati, J. C. & Salam, A. *Phys. Rev. Lett.* **31**, 661 (1973); *Phys. Rev.* **D8** 1240 (1973); **D10**, 275 (1974).
6. Georgi, H. *Nature* **288**, 649 (1980).
7. Glashow, S. L. & Nanopoulos, D. V. *Nature* **281**, 464 (1979).
8. Farhi, E. & Susskind, L. Preprint TH. 2975 CERN (1980).
9. Ellis, J. Preprint TH. 2942 CERN (1980).
10. Learned, J., Reines, F. & Soni, A. *Phys. Rev. Lett.* **43**, 907 (1979).
11. Ellis, J., Gaillard, M. K., Maiani, L. & Zumino, B. *Unification of the Fundamental Particle Interactions* (eds Ferrara, S., Ellis, J. & van Nieuwenhuizen, P.) 69 (Plenum, New York 1980).
12. Ellis, J., Gaillard, M. K. & Zumino, B. *Phys. Lett.* **94B**, 343 (1980).
13. Georgi, H., Quinn, H., R. & S. Weinberg, *Phys. Rev. Lett.* **33**, 451 (1974).
14. Stueckelberg, E. C. G. & Peterman, A. *Helv. phys. Acta* **26**, 499 (1953).
15. Gell-Mann, M. & Low, F. E. *Phys. Rev.* **95**, 1300 (1954).
16. Buras, A. J., Ellis, J., Gaillard, M. K. & Nanopoulos, D. V. *Nucl. Phys.* **B135**, 66 (1978).
17. Marciano, W. J. & Sirlin, A. *Phys. Rev. Lett.* **46**, 163 (1981).
18. Barber, D. P. *et al.* *Phys. Lett.* **89B**, 139 (1979).
19. Bartel, W. *et al.* *Phys. Lett.* **91B**, 142 (1980).
20. Brandelik, R. *et al.* *Phys. Lett.* **94B**, 437 (1980).
21. Berger, C. *et al.* *Phys. Lett.* **97B**, 459 (1980).
22. Ellis, J., Gaillard, M. K., Nanopoulos, D. V. & Rudaz, S. *Nucl. Phys.* **B176**, 61 (1980).

Dynamical symmetry in the quadratic Zeeman effect

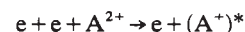
C. W. Clark & K. T. Taylor

Science Research Council, Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK

Zimmerman, Kash and Kleppner¹ have recently posulated the existence of a 'hidden' symmetry in atomic hydrogen in a uniform magnetic field. We have drawn similar conclusions from independent work². Here, we show that the quasi-constant of the motion in this system can be related in a simple way to the asymptotic form of the quantum mechanical wavefunction. This development is consistent with the broad outline drawn by Fano^{3,4} who has emphasized the role of this particular problem as a member of a large and important class of dynamical systems characterized by the existence of motion along a 'ridge' of a potential surface.

Such motions are intrinsically unstable, because a fall off the ridge into a region of lesser potential is always possible. Nevertheless, it is frequently the case that important observable

quantities are determined solely by these unstable motions. This has long been known^{5,6} to be so for the most prominent other member of this class, the system of two electrons escaping simultaneously from the field of an ion, which occurs in problems of electron impact ionization and multiple photoionization of atoms. Thus, a thorough understanding of the quadratic Zeeman effect, necessary in itself for the interpretation of current experiments, may also be relevant to more general questions of atomic dynamics. We draw special attention here to the analogy provided by the problem of two-electron escape, in which competing modes of partial escape



give rise to the Wannier threshold law^{5,6}.

We focus our discussion on the interpretation of Fig. 1, the bottom frame of which shows the oscillator strength for dipole transitions from the ground state of hydrogen to excited states with a single quantum of angular momentum in the direction of the magnetic field. These 'experimental' data were obtained by diagonalizing the hamiltonian within a large basis ($\sim 1,500$) of sturmian functions, using a set of banded matrix routines developed at the National Physical Laboratory which have been implemented on the CRAY-1 computer at Daresbury. The field strength is 4.7 T. Transitions to these states would result from absorption of light with circular polarization about the axis of the magnetic field. The lowest cluster of lines shown corresponds to the perturbed manifold of hydrogenic states with odd parity and principal quantum number $n = 23$. As the energy of the final

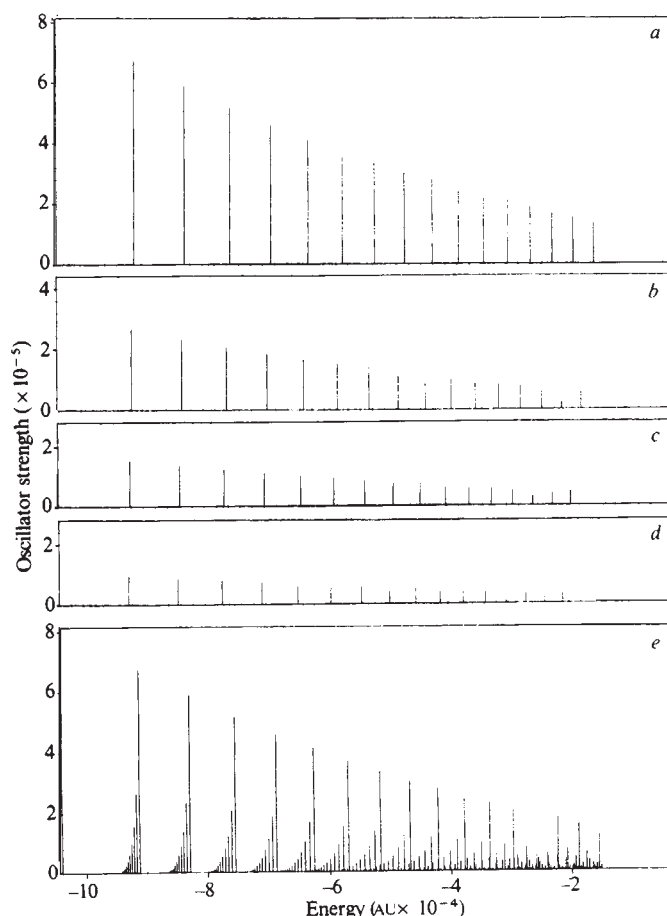


Fig. 1 The oscillator strength for absorption of circularly polarized light by atomic hydrogen in the ground state, for photon energies below the ionization limit. a, The principal series; b-d, second, third and fourth series; e, all lines present in this energy range. The abscissa is the total energy minus the linear Zeeman shift.

state increases the line clusters interpenetrate, with a surprising lack of mutual perturbation among lines of comparable oscillator strength with nearly equal energy. We are thus led to view this spectrum as a superposition of many essentially independent line series. The first four such series, in order of spectroscopic prominence, are drawn out in the upper frames of Fig. 1 (the 16 lowest members of each series only are displayed in these frames). Complementary evidence for this classification scheme can be inferred from the independent work of Zimmerman *et al.*¹, who have shown the dependence of lower energy levels on magnetic field strength at stronger fields. Their plots exhibit many near crossings of energy levels, which would be forbidden on general grounds⁷ except in the presence of an additional constant of the motion.

Corroborative information is obtained by examining the appropriate solutions to the quantum mechanical equations of motion. The hamiltonian for this system, if the centre-of-mass motion and relativistic effects are disregarded, is the same as that for a hydrogen atom in the absence of a field, with an additional term describing the interaction of the electron with the field. This term contains a part that is linear in the field strength and depends only on the projection of the electron's angular momentum on the field direction; this gives rise to the familiar linear Zeeman effect, which consists of a constant energy shift of all lines in the spectrum of Fig. 1. The remainder of the magnetic interaction is quadratic in the field strength, and results in the Schrödinger equation (in atomic units)

$$\left(-\frac{1}{2} \frac{\partial^2}{\partial r^2} + \frac{\vec{l}^2}{2r^2} - \frac{1}{r} + \frac{1}{2} \beta^2 r^2 \sin^2 \theta \right) \psi = \varepsilon \psi \quad (1)$$

where $\vec{l}$ is the electron angular momentum operator, β one-half the magnetic cyclotron frequency, θ the angle between the magnetic field $\vec{B} = B\hat{z}$ and the electron position vector $\vec{r}$, and ε is the total energy minus the linear Zeeman shift. The last term on the left-hand side of equation (1) is a repulsive potential which assumes its maximum value in the plane $z = 0$. Because of the very small value of β ($\beta = 10^{-5}$ for a 4.7 T field), this term has a significant effect on the motion only at large distances r , where the potential surface takes the form indicated in Fig. 2. The equipotential surfaces are essentially spherical at small r , and cylindrical at large r . Thus, in photoabsorption the initial state of the electron, being tightly bound to the nucleus, is essentially spherical; the final state, which has significant amplitude at large distances, must eventually conform to the cylindrical geometry of the potential surface there. Understanding this transition between disparate geometries has constituted a major difficulty for the theory. In the problem of double photoionization also, the initial state has a more or less spherical symmetry, and the two electrons are not too strongly correlated. Dynamical analysis shows, however, that the only final state configurations which

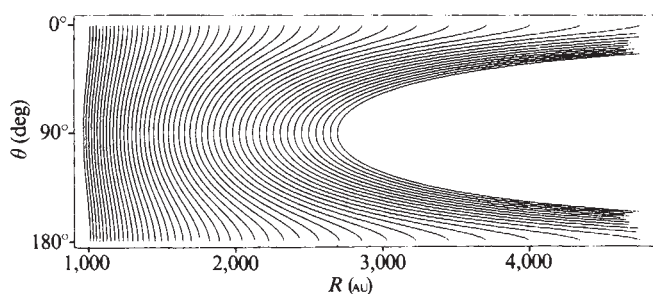


Fig. 2 The equipotential lines for electron motion in the field of a proton and a 4.7 T magnetic field. The ordinate is the angle θ subtended by the position vector and the magnetic field, the abscissa the electron-proton distance, R , in units of Bohr radius. The rightmost equipotential line corresponds to the energy $\varepsilon = 0$; the energy spacing between successive lines is one-fifth of the magnetic cyclotron frequency. A potential ridge is apparent at the angle $\theta = 90^\circ$.

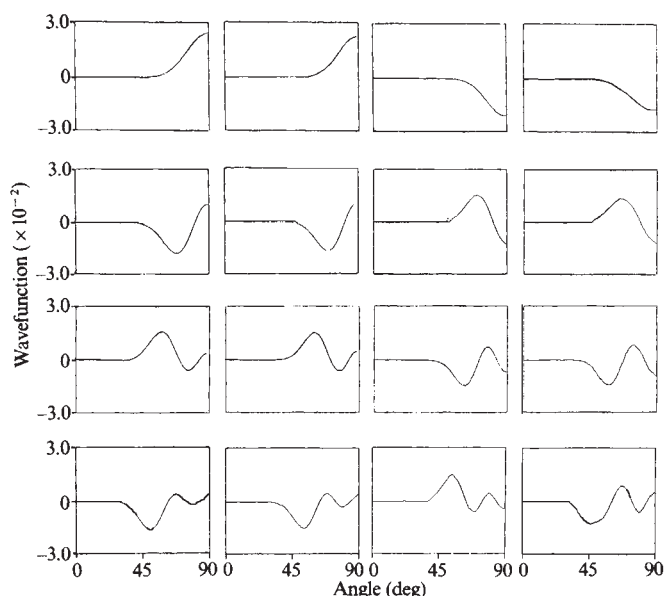


Fig. 3 Wavefunction versus angle along the appropriate turning line for selected states of the first four series. First row: the 1st, 5th, 10th and 16th states of the principal series as indicated in Fig. 1; second and third rows: the corresponding states of the second and third series; fourth row: the 1st, 5th, 10th and 15th states of the fourth series, the 16th state being strongly perturbed. All wavefunctions are symmetric about 90° .

lead to double ionization are those in which the two electrons lie at equal and opposite distances from the ion. Such configurations are unstable and can be considered as located on a ridge of the potential describing the interaction among the three particles of the system. A fall off the ridge results in capture of one of the electrons into a high Rydberg state. In our Zeeman problem the quadratic potential displays a ridge in the plane $z = 0$. We shall show that a large fraction of the oscillator strength is associated with motion of the electron along this ridge, as has been suggested by early semi-classical treatments⁸ of the problem. In this case the fall off the ridge leads to escape of the electron along the magnetic field.

We consider first the principal series of Fig. 1, the members of which are shown in the top frame. We have computed accurate wavefunctions for the states of this series using the method indicated above, which has been described in more detail elsewhere². Briefly, each wavefunction is represented by expansion in terms of products of spherical harmonics and sturmian functions, the coefficients of which are determined by the diagonalization of the hamiltonian matrix. In the absence of a magnetic field these states would reduce to hydrogen wavefunctions with angular momentum $l = 1$, and would carry all the oscillator strength. If the effect of the field is regarded as perturbative, which is appropriate for the lowest members, some higher angular momenta are mixed in with a resulting diminution in the oscillator strength. The higher members of this series are not characterized by the predominance of any one value of the angular momentum at small r . If, however, we look at the wavefunction at large r a simple consistent pattern emerges. Figure 3 shows the variation of the wavefunctions for some of these states along their classical turning lines, that is, those lines of Fig. 2 for which the equipotential energy is equal to the energy of the states. The wavefunctions are evaluated simply by summing the weighted individual values of the sturmian and spherical harmonic functions, computed by recursion formulae over the locus of the equation $\frac{1}{2} \beta^2 r^2 \sin^2 \theta - 1/r = \varepsilon$. It is seen that the states of the principal series may be characterized by having no nodes along their turning lines, and have a significant concentration of amplitude on the ridge. If, on the other hand, one examines the wavefunctions for these states in the $z = 0$ plane, it will be seen that each successive state has an additional

node along the ridge. This is the source of the rather close agreement of the energies of these levels with that predicted by the simple two-dimensional WKB theory^{2,8}. The subsequent series admit a similar classification. The states of the second series exhibit two nodes along the turning line, those of the third exhibit four, and those of the fourth, six—the number of nodes increases by two between series because all these states are necessarily even under reflection in the plane. The states with odd numbers of nodes, not shown in Fig. 1, have wavefunctions which vanish on the ridge and have much smaller oscillator strengths; these would be obtained in absorption of light polarized in the direction of the magnetic field.

Thus, the magnitude of the oscillator strength associated with a given state is directly correlated with the localization of its wavefunction on the ridge at large distances. Moreover, the spread of the wavefunction off the ridge occurs in a regular manner for the states of greatest spectral prominence. The members of a given series differ by their degree of excitation along the rising ridge of the potential; the separate series are distinguished by the degree of excitation of the unstable motion across the ridge. This picture of two nearly uncoupled modes of oscillation, which holds at least throughout the energy range of Fig. 1, explains the observed weakness of interseries perturbations. Note that this classification is an approximate one, and that some states, for example the 15th member of the third series and the 16th member of the fourth, show marked departures from the regular scheme. However, these isolated irregularities occur, in every case examined so far, when lines from separate series coalesce. Although one cannot extrapolate these results to higher energies or other atoms with complete confidence, nevertheless, it seems probable that the major structures seen in photoabsorption cross-sections at energies above the ionization limit⁹ are due to resonance states which can be placed in a principal series. The secondary features which are seen would then be associated with the higher series, and should be expected to be broadened more rapidly with increasing energy due to the greater probability of escape off the ridge.

Localization of a quantum mechanical wavefunction in a region of maximum potential, although at first seemingly contradictory, is probably a phenomenon of general importance. For example, the occurrence of barrier top resonances in heavy ion direct reactions have been analysed by Friedman and Goebel¹⁰. Some qualitative remarks are suggested by the form of the line clusters in the low energy region of Fig. 1e. If viewed in the framework of degenerate perturbation theory, the energy levels within a cluster are isomorphic to the eigenfrequencies of a chain of oscillators with nearest-neighbour coupling¹¹. If the spring constants along such a chain were uniform, a sinusoidal distribution of energy levels would be obtained and the eigenmodes of oscillation would be standing waves. This occurs, for instance, in the simple Hückel model for a π electron system of conjugated double bonds, resulting in electron wavefunctions which are entirely delocalized. If, however, the coupling along the chain is non-uniform—corresponding to the insertion of a substitutional impurity in a Hückel molecule—the eigenmodes of oscillation must become partially localized. In our problem the analogous nearest-neighbour coupling is non-uniform, and gives rise to a linear distribution of energy levels at the extremities of each cluster. The wavefunctions of states at the top and the bottom of a cluster are then necessarily localized, respectively, in the regions of maximum and minimum potential.

Received 6 March; accepted 16 June 1981.

1. Zimmerman, M. L., Kash, M. M. & Kleppner, D. *Phys. Rev. Lett.* **45**, 1092–1094 (1980).
2. Clark, C. W. & Taylor, K. T. *J. Phys.* **B13**, L737–L743 (1980).
3. Fano, U. *Phys. Rev.* **A22**, 2660 (1980).
4. Fano, U. *J. Phys.* **B13**, L519 (1980).
5. Wannier, G. H. *Phys. Rev.* **90**, 817–825 (1953).
6. Rau, A. R. P. *Phys. Rev.* **A4**, 207–220 (1971).
7. Von Neumann, J. & Wigner, E. *Phys. Z.* **30**, 467 (1929).
8. Starace, A. F. *J. Phys.* **B6**, 585–590 (1973).
9. Lu, K. T., Tomkins, F. S. & Garton, W. R. S. *Proc. R. Soc. A* **362**, 421–424 (1978).
10. Friedman, W. A. & Goebel, C. J. *Ann. Phys.* **104**, 145–183 (1977).
11. Clark, C. W. *Phys. Rev.* **A24** (in the press).

Failure of Arrhenius equation for hydroxyl radical–bicarbonate ion reaction above 100 °C

David Ritchie McCracken*

CEGB, Berkeley Nuclear Laboratories, Berkeley, Gloucestershire GL13 9PB, UK

George V. Buxton

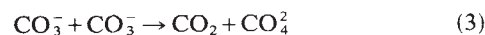
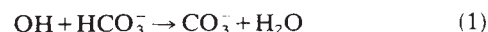
The University of Leeds, Cookridge Radiation Research Centre, Cookridge Hospital, Leeds LS16 6QB, UK

Although the chemical kinetics of reactive species in water at temperatures above 100 °C has received very little attention, the radiation chemistry of aqueous coolant at the high pressures (150 bar) and temperatures (330 °C) of pressurized and boiling-water nuclear reactors is of obvious interest to reactor chemists, engineers and designers¹. There is a wealth of data² on rate constants for reactions of the radiolytically important radicals e_{aq}^- , H, OH, O⁻, HO₂ and O₂⁻ at room temperature, but only a few rates have been measured³ up to 90 °C and only one⁴ up to 200 °C. It has thus become common^{5–9} to calculate the rate constants required for modelling water radiolysis in reactor conditions by extrapolating room temperature data using the Arrhenius equation $k = A \exp(-E/RT)$ where E , the assumed activation energy, has a value in the range 12–20 kJ mol⁻¹. We report here the failure of this method to predict the rate of reaction of hydroxyl radicals with bicarbonate at temperatures up to 200 °C.

We are making a systematic study of the reaction kinetics of the primary radicals and their secondary products, formed in the radiolysis of water, using pulse radiolysis–kinetic spectrophotometry over the temperature range 25–200 °C. Simultaneously, measurements are being made of the effect of temperature on yields of the primary radicals, as this is another area where data are sparse and conflicting^{7,9–12}.

We began by studying the hydroxyl radical reaction with bicarbonate ion for two reasons: first, this reaction has already been investigated at room temperature¹³, and second, it will occur when alkaline water is exposed to air and radiation, as happens in nuclear power operations.

The rate of reaction (1) was measured by observing the formation of the carbonate radical ion CO₃⁻ which has an optical adsorption band with a maximum at 600 nm, in nitrous oxide-saturated solutions of sodium bicarbonate at natural pH.



Conditions were chosen to ensure that reaction (2) was too fast, and reaction (3) too slow, to interfere in the measurement of k_1 .

The solution was contained in a conventional Pyrex–quartz flow system inside a pressure vessel and subjected to an overpressure of helium at 20 bar to prevent boiling. The temperature of the solution was measured with a thermocouple in a quartz sleeve which protruded into the radiolysis cell. The thermocouple was checked by observing the boiling point of water as a function of pressure and this showed that the error in the temperature readings was ≤ 2 °C. The radiation source was a Van de Graaff accelerator producing 0.6- μ s pulses of 3-MeV

* Present address: Atomic Energy of Canada, Chalk River Nuclear Laboratories, Ontario, Canada.

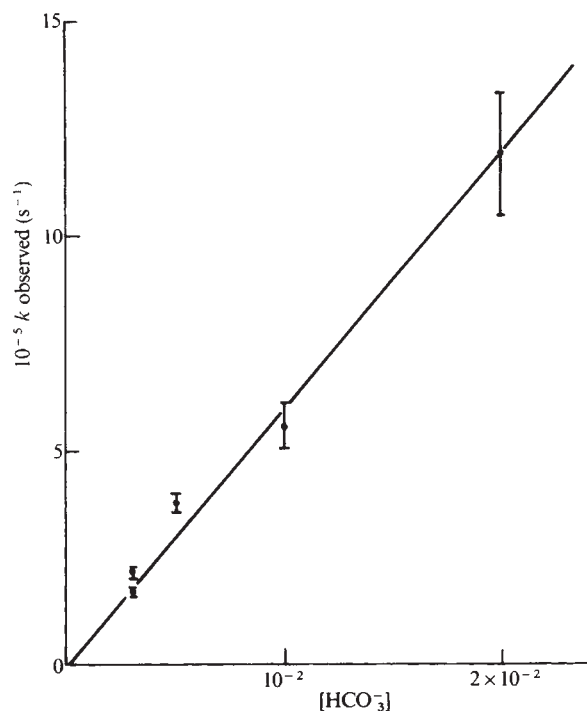


Fig. 1 Observed first-order rate constant for build up of CO_3^- absorption at 308 K.

electrons. The dose per pulse was ~ 1 krad and five pulses could be given to each sample of solution without producing a detectable change in the results. The sample was renewed at each temperature.

There was no change in either the shape or the intensity of the spectrum of CO_3^- with temperature. As $G(\text{CO}_3^-) = G(e_{\text{aq}}^-) + G(\text{OH})$, this indicates that there is little change in the yield of these primary species up to 200 °C.

The dependence of the observed pseudo-first-order rate constant on $[\text{HCO}_3^-]$ is shown in Figs 1 and 2 for 35 and 196 °C respectively. Second-order rate constants are plotted according to the Arrhenius equation in Fig. 3. Here the data have been corrected for changes in activity resulting from changes in the density of water¹⁴ and the equilibrium constants¹⁵ K_4 and K_5 .



Activity coefficients were calculated using extended Debye-Hückel theory¹⁶ with appropriate constants¹⁷ for the temperature range 25–200 °C. The largest correction (for $3 \times$

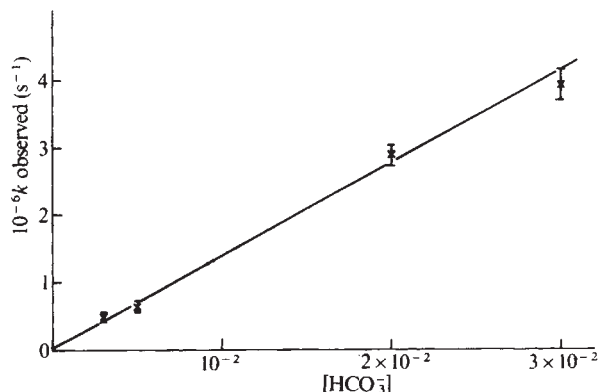
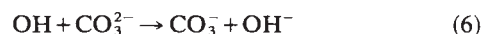


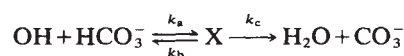
Fig. 2 Observed first-order rate constant for build up of CO_3^- absorption at 469 K not corrected for changing density or activity, with temperature. True $k = k_{\text{obs}}/0.84$.

10^{-3} mol dm^{-3} bicarbonate at 200 °C) was only 5%, and the most important factor was the density of water, for which the largest correction was 14%. In all cases $[\text{CO}_3^{2-}] \leq 0.01 [\text{HCO}_3^-]$ and we estimate the contribution of reaction (6) to the observed rate of formation of CO_3^- to be $(10 \pm 1)\%$ over the whole range of temperature and $[\text{HCO}_3^-]$ (based on our unpublished measurements of $k(\text{OH} + \text{CO}_3^{2-})$ in the range 25–200 °C). Accordingly, the data in Fig. 3 have been reduced by 10%, although such a correction does not alter the shape of the Arrhenius plot and is less than the spread of the data ($\pm 20\%$) about the solid curve.



There is no evidence for reaction of OH with carbonic acid and we assume that it can be neglected.

Figure 3 clearly shows that the data for k_1 do not obey the Arrhenius equation, but the deviation is not obvious below 100 °C. The low temperature data extrapolate to $k_1 \approx 10^9 \text{ dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$ at 330 °C with an apparent activation energy $\sim 14 \text{ kJ mol}^{-1}$, whereas the high temperature results give $k_1 \approx 2 \times 10^8 \text{ dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$ and show scarcely any apparent activation energy. A reaction scheme which is consistent with the data is



where X is an adduct of OH and HCO_3^- . The adduct could be a

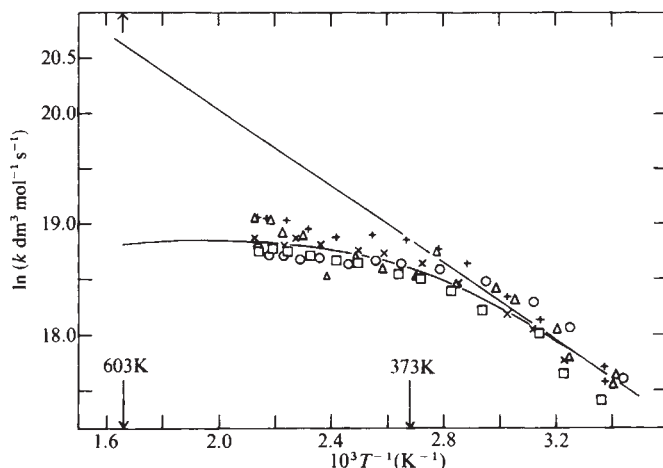
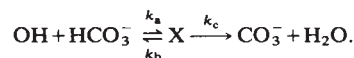
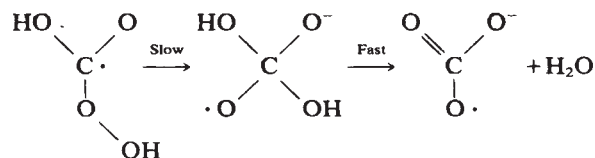


Fig. 3 $\ln$ (corrected observed second-order rate constant) as a function of reciprocal temperature (K^{-1}). The solid curve is a fit of the data according to the mechanism



$\square$, 3×10^{-2} ; $\times$, 2×10^{-2} ; $+$, 1×10^{-2} ; $\circ$, 5×10^{-3} ; $\triangle$, 3×10^{-3} mol dm^{-3} HCO_3^- .

carbon-centred radical, which rearranges to an oxygen-centred one and then splits off water to leave CO_3^- .



Then,

$$k_{\text{obs}} = \frac{k_a k_c}{k_b + k_c} = \frac{A_a \exp(-E_a/RT)}{1 + (A_b/A_c) \exp(E_c - E_b/RT)}$$

The solid curve in Fig. 3 has been calculated with $A_a = 1.3 \times 10^{11} \text{ dm}^3 \text{ mol}^{-1} \text{ s}^{-1}$; $A_b/A_c = 1.2 \times 10^3$; $E_a = 18.9 \text{ kJ mol}^{-1}$; $E_b - E_c = 21.3 \text{ kJ mol}^{-1}$. Unfortunately, the data

have not been obtained over a sufficiently large temperature range to fix these parameters to better than 50% for E_a and $E_b - E_c$ or within an order of magnitude for A_a and A_b/A_c . Experiments are planned which will extend the temperature range to $\geq 330^\circ\text{C}$.

In spite of the present limitations in the data, they demonstrate quite clearly that the extrapolation of linear Arrhenius plots of rate constants measured at low temperatures ($< 100^\circ\text{C}$) to obtain their high temperature values could be quite misleading, and that experimental measurements at temperatures pertaining to nuclear reactor operations need to be made if water radiolysis is to be modelled correctly in these conditions.

We thank Rod E. Brown, Brian Hale, Frank Hambleton and Philip Wright for technical assistance and advice.

Received 6 April; accepted 10 June 1981.

1. Cohen, P. *Water Coolant Technology of Power Reactors* (Gordon and Breach, Edinburgh 1969).
2. *Selected Specific Rates of Reactions of Transients from Water in Aqueous Solution* (NSRDS-NBS 43 (1973); 51 (1975); 59 (1977)).
3. Schmidt, K. H. *J. phys. Chem.* **81**, 1257 (1977).
4. Christensen, H. & Sehested, K. *Radiat. phys. Chem.* **16**, 183 (1980).
5. Jenks, G. H. *ORNL-3848* (1965).
6. Turner, S. E. *Design Construct. Pract.* **12**, 66 (1968).
7. Michael, B. D., Hart, E. J. & Schmidt, K. H. *J. phys. Chem.* **75**, 2798 (1971).
8. Burns, W. G. & Moore, P. B. *Radiat. Eff.* **30**, 233 (1976).
9. Burns, W. G. & Marsh, W. R. *JCS Faraday I* **77**, 197 (1981).
10. Jha, K. N., Ryan, T. G. & Freeman, G. R. *J. phys. Chem.* **79**, 868 (1975).
11. Kalecinski, J. *Bull. Acad. Pol. Sci.* **21**, 209 (1973).
12. Burns, W. G., Marsh, W. R. & Barton, R. A. *Nature* **241**, 86 (1973).
13. Buxton, G. V. *Trans. Faraday Soc.* **65**, 2150 (1969).
14. Engineering Sciences Data Unit, No. 68010 (1978).
15. *Chem. Soc. Spec. Publ.* No. 25 (1971).
16. Moore, W. J. *Physical Chemistry* (Longmans, New York, 1962).
17. Eisenberg, D. & Kauzmann, W. *The Structure and Properties of Water* (Oxford University Press, 1969).

Radio-echo layering in polar ice sheets and past volcanic activity

D. H. M. Millar

Scott Polar Research Institute, University of Cambridge,
Lensfield Road, Cambridge CB2 1ER, UK

A recent study¹ has shown how acidity profiles along dated Greenland ice cores can reveal large volcanic eruptions in the Northern Hemisphere during the past 10,000 yr. It has been suggested that these layers of acidic ice in the polar ice sheets may also be detected by airborne radio-echo sounding (RES) techniques as stratification echoes²⁻⁴. Explosive volcanic eruptions eject large amounts of SO_2 into the stratosphere⁵ where it forms an H_2SO_4 aerosol⁶. Studies of Antarctic snow⁷ and Greenland ice cores² show that this material can be deposited in layers of large areal extent on polar ice sheets. Calculations based on observations of these slightly acidic layers show that they should give rise to radar reflections of similar magnitude to those observed for stratification echoes. New RES data from the Antarctic enables the present comparison to be made between observed layer power reflection coefficients (PRCs), and calculated values for reflections from acidic ice layers and from layers of ice of changed density. A gap in layering has been identified which seems to coincide with a similar gap reported in Greenland at Crête⁸; profiles of layer PRC against age show a common pattern for many sites on the Antarctic ice sheet. This PRC/age profile may provide a record of explosive volcanic activity for the Southern Hemisphere over the past 150,000 yr.

During a joint Antarctic RES Programme conducted by the NSF, the Scott Polar Research Institute and the Electromagnetics Institute of the Technical University of Denmark (TUD), calibrated A-scope radar displays (showing returned/transmitted power versus range, Fig. 1) from the 60 MHz TUD ice-sounding radar were recorded, from which PRCs for internal layers were obtained after allowing for geometrical

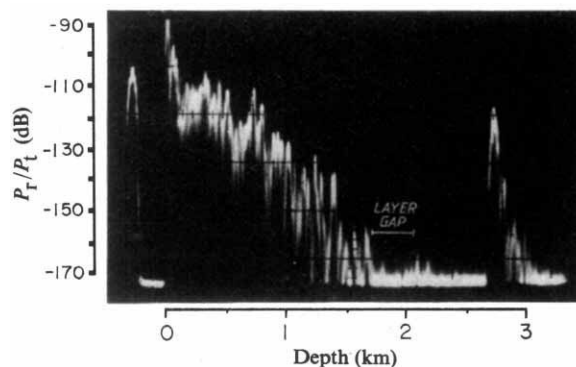


Fig. 1 Calibrated A-scope record from near Crête, Greenland. Note gap in layering at 1,700–2,100-m depth.

spreading and absorption losses⁹ to give profiles of layer PRC against depth (see Fig. 2a). Errors are estimated to be ± 10 m for depths and ± 2 dB for PRCs; layer roughness causes the echo strength to fluctuate rapidly over a 10 dB range as the aircraft moves so that repeated, closely-spaced measurements are necessary to average out this effect.

Fluctuations in ice density have been suggested⁹⁻¹¹ as a cause of layer echoes in ice sheets, and stratigraphic analysis¹² of ice cores indicates that clear ice layers (millimetres thick) and similar features do occur in the upper parts. A simple expression for the maximum density contrast between clear ice and 'bubbly' polar ice as a function of depth⁹ (which compares well with estimates from the density profile of Byrd Station, Antarctica) has been used to compute the upper limit of PRCs from this mechanism (Fig. 2a, solid line). At Byrd observed PRCs are ~ 10 –20 dB weaker than this limit, but decrease with depth in the same way; this is consistent with the observation of occasional density fluctuations in the core¹⁰.

At Byrd the ice is completely bubble-free below 1,100 m (ref. 13) and so this mechanism cannot apply to the deeper echoes seen. Paren and Robin¹⁴ suggested that deeper layer echoes are caused by layers of ice of changed loss tangent, due to changed impurity levels, and Hammer² thought that acids of volcanic origin may provide the impurity. It is suggested here that varying concentrations of such acids are the only plausible cause for changes in loss tangent detectable by RES. The highest acidity ice layers observed in the 400 m core from Crête are believed⁴ to coincide in depth with layer echoes, however, no comparison of reflection coefficients was attempted. In the absence of experimental measurements of the loss tangent of acidic polar ice, approximate estimates of the PRC for a layer of ice containing an elevated H^+ concentration are made from laboratory measurements¹⁵ of the high frequency (h.f.) conductivity of HF-doped ice. The difference in loss tangent, $\Delta(\tan \delta)$, between pure ice and ice doped with acid is in theory¹⁶ related to the associated change in h.f. conductivity, $\Delta\sigma_\infty$:

$$\Delta(\tan \delta) = \Delta\sigma_\infty / 2\pi f \epsilon' \epsilon_0 \quad (1)$$

where ϵ' is the permittivity of polar ice, ϵ_0 the permittivity of

Table 1 Observed concentrations of acids in polar ice sheets following known volcanic eruptions, and estimated high frequency conductivities and power reflection coefficients

Eruption	Peak $[\text{H}^+]$ (mmol m^{-3})	Peak h.f. conductivity ($\text{Sm}^{-1} \times 10^{-6}$)	R_s (dB)
Thera (1390 BC)	4.5	7.0	-83
Eldgja (AD 934)	20.0	30.0	-64
Hekla (1104)	5.5	9.0	-79
Laki (1783)	21.0	30.0	-64
Tambora (1815)	8.0	15.0	-72
Katmai (1912)	4.5	7.0	-83

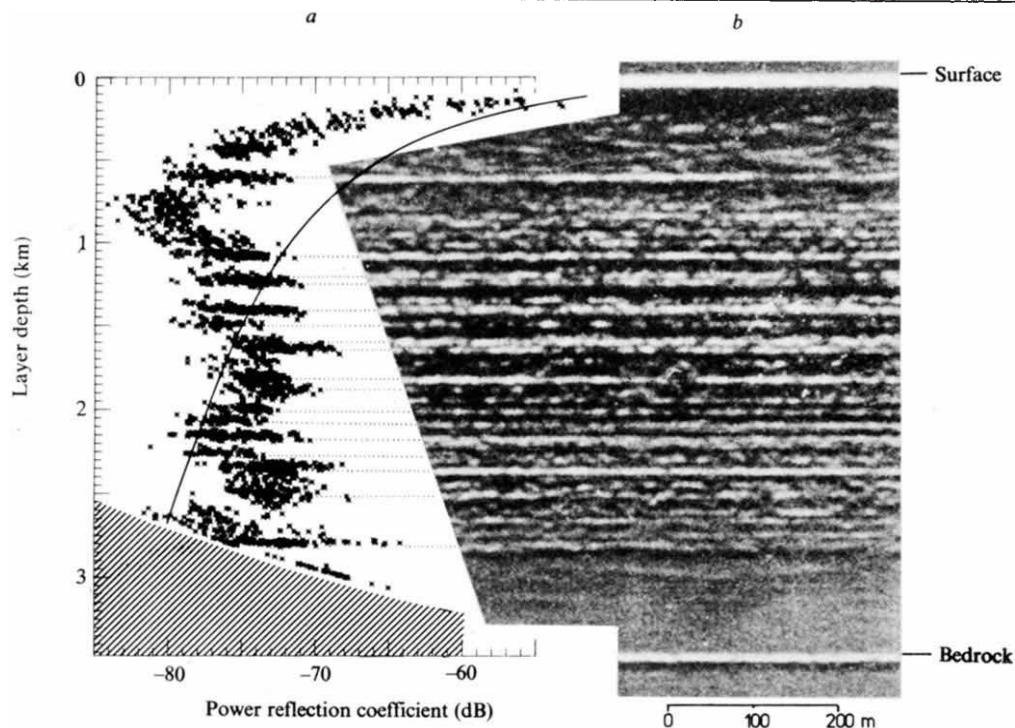


Fig. 2 *a*, PRC/depth profile (at site ~100 km north-west of Dome C) from many power measurements obtained with the TUD 60 MHz radar, using a 250-ns pulse length, 4-MHz receiver bandwidth and ~10-kW peak power. Solid line shows maximum PRC expected from density fluctuations, and shaded area is that beyond the theoretical detection limit. *b*, Corresponding section of Z-scope film (a differentiated, intensity-modulated version of the power return against time along flight-track); dotted lines connect groups of PRC measurements with layers.

free space and f the radar carrier frequency. The high mobility of the H^+ ion is thought to dominate the h.f. dielectric properties of acid-doped ice, and so the nature of the anion will be expected to be of secondary importance. Impurities other than H^+ do not alter the loss tangent sufficiently (except in very high concentrations) to give detectable echoes. PRCs for a single interface (R_s) and a thin layer (R) of ice of changed loss tangent may then be obtained from the relations given by Paren and Robin:

$$R = 4 \sin^2(2\pi l/\lambda) R_s \quad (2)$$

$$R_s = 1/16[\Delta(\tan \delta)]^2 \quad (3)$$

where l is the layer thickness and λ the radar wavelength in ice. Table 1 shows measured H^+ concentrations associated with several large historical eruptions³, and estimated h.f. conductivities and PRCs; errors in estimating R_s are thought to be <5 dB. The estimated values are similar to those observed in the deeper ice below the region where density effects dominate. Little temperature dependence of PRC on H^+ concentration is estimated at temperatures above -40°C , so any variations of PRC with depth are expected to reflect past volcanic acid input to the ice sheet. PRC/layer age profiles from widely spaced Antarctic sites show many common features (Fig. 3), and this pattern may thus be interpreted as a record of the varying amounts of volcanic acid deposited on the ice sheet at different times.

Dated Z-scope layering from several Antarctic sites also reveals a common gap at ~15–20 kyr BP (this is not obvious on individual Z-scope records due to the close spacing of layers in the Antarctic), which seems to be simultaneous with a similar gap in layering at 1,900–2,300 m depth at Crête⁸ and dated¹⁷ at ~13–23 kyr BP. The uniqueness and coincidence of these gaps suggest a common cause: one possibility is chemical neutralization of volcanic acids by calcareous dust blown on to the ice sheets at the end of the last glaciation, due to exposure of former shallow water areas. Large increases in ice calcium levels are observed at Camp Century (Greenland) and Byrd at this time¹⁸, and at Camp Century the ice is known¹ to be slightly alkaline during this period. Such neutralization could affect the PRC/age profiles at other times, but this is thought to be unlikely because of the steadier Ca levels¹⁸ before 20 kyr BP, the probable link between Ca dust production and the end of the glaciation, and the lack of any other gaps in Z-scope layering. The PRC/age

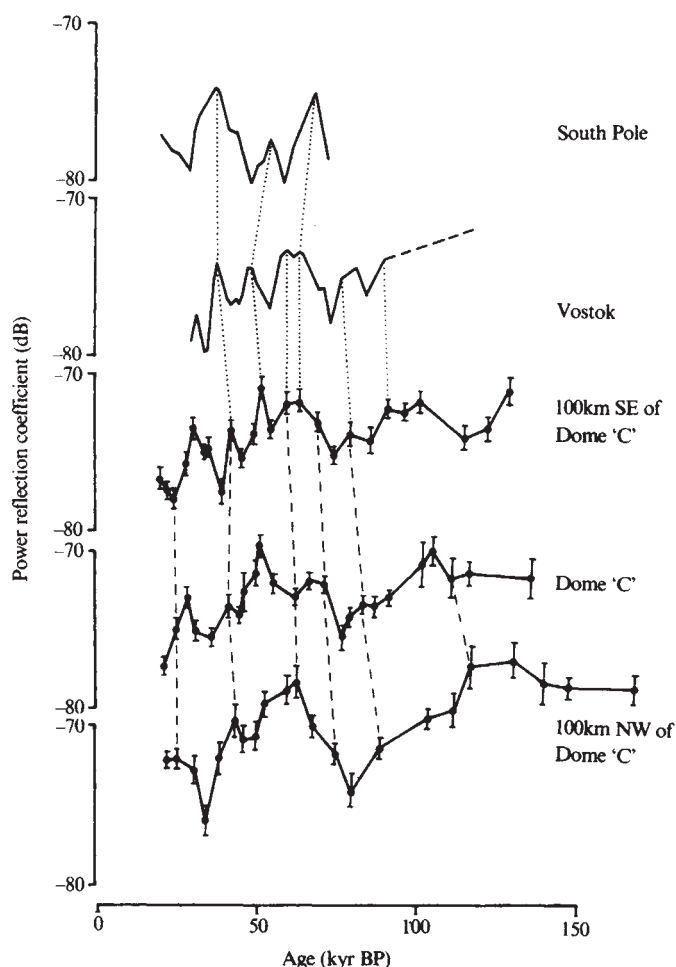


Fig. 3 PRC/age profiles for layers observed at South Pole, Vostok, Dome C and two sites (1) ~100 km south-east and (2) ~100 km north-west of Dome C. Bars indicate standard errors in power reflection coefficients and dashed lines join layers followed between sites on the Z-scope record. Dotted lines indicate suggested common features (data do not exist to enable layers to be followed on the Z-scope record).

pattern may therefore provide a useful indication of past volcanic acid input to the ice sheets, reflecting explosive volcanic activity in the Southern Hemisphere.

Conventional airborne echo sounding has already proved valuable in revealing the flow structure of the ice. The extension of the technique, using purpose-built radars (with better resolution) to the field of ice chemistry offers advantages of speed, economy, and the ability to sample the ice over representative areas, rather than at isolated points (with inherent uncertainties due to accumulation variability) plus deep penetration and consequently a long time scale. Such a technique should be complementary to drilling as core results are necessary to 'calibrate' the radar results.

I thank Dr G. de Q. Robin for supervision, and Drs C. U. Hammer, C. S. Neal, S. Overgaard and J. G. Parn for useful discussions and encouragement. Dr S. Overgaard (TUD) gave permission to reproduce Fig. 1. Logistic support was provided by the NSF and financial support in the form of a Research Studentship from the NERC was much appreciated.

Received 8 February; accepted 10 June 1981.

1. Hammer, C. U., Clausen, H. B. & Dansgaard, W. *Nature* **288**, 230–235 (1980).
2. Hammer, C. U. *Nature* **270**, 482–486 (1977).
3. Hammer, C. U. *J. Glaciol.* **25**, 359–372 (1980).
4. Gudmandsen, P. & Overgaard, S. *Technical University of Denmark, Electromagnetics Institute, Rep. P312* (1978).
5. Lazrus, A. L. *et al. J. geophys. Res.* **84**, 7869–7875 (1979).
6. Arnold, F. & Fabian, R. *Nature* **283**, 55–57 (1980).
7. Delmas, R. & Bouton, C. *J. geophys. Res.* **85**, 5645–5649 (1980).
8. Gudmandsen, P. *J. Glaciol.* **15**, 95–101 (1975).
9. Robin, G. de Q., Evans, S. & Bailey, J. T. *Phil. Trans. R. Soc. A* **265**, 437–505 (1969).
10. Clough, J. W. *J. Glaciol.* **18**, 3–14 (1977).
11. Ackley, S. F. & Keliher, T. E. *J. Geophys. Res.* **84**, 5675–5680 (1979).
12. Langway, C. C. *Jr Geol. Soc. Am. Spec. Pap.* **125** (1970).
13. Gow, A. J. *Int. Symp. Antarctic Glaciological Exploration, Hanover 78–90* (1970).
14. Parn, J. G. & Robin, G. de Q. *J. Glaciol.* **14**, 251–259 (1975).
15. Camplin, G. C., Glen, J. W. & Parn, J. G. *J. Glaciol.* **21**, 123–142 (1978).
16. Hobbs, P. V. *Ice Physics* (Clarendon, Oxford, 1974).
17. Hammer, C. U. *et al. J. Glaciol.* **20**, 3–26 (1978).
18. Cragin, J. H., Herron, M. M., Langway, C. C. & Klouda, G. *Proc. Polar Oceans Conf. Montreal*, 617–631 (1977).

Intermediate depth seismicity in the western Mediterranean unrelated to subduction of oceanic lithosphere

D. Hatzfeld

Laboratoire de Geophysique Interne, IRIGM, Université Scientifique et Médicale de Grenoble, BP 53X, 38041 Grenoble Cedex, France

M. Frogneux

Service de Physique du Globe, Institut Scientifique, av. Ibn Batouta, Rabat, Morocco

There are three zones in the Gibraltar area where intermediate depth earthquakes occur down to 150 km. Two of these, the Gulf of Cadiz and the western Alboran Sea, are oceanic areas which could be associated with the subduction of oceanic lithosphere. Our present measurements of arrival times of P and S waves show, however, that the third zone beneath the High Atlas is a typical intracontinental chain, where it is not likely that oceanic lithosphere has been subducted during the past 20 million years.

In the oceanic domain, West of 15°W, the Azores–Gibraltar seismicity clearly outlines the Eurasian–African plate boundary. Towards the East, however, where the continental domain begins, previous studies have shown the presence of widely scattered epicentres^{1–3} that have led to postulations of an Alboran buffer sub-plate. Precise relocations reduced the scatter and the seismicity can rather be interpreted as a reactivation of old tectonic faults^{4,5}.

Table 1 Velocity structure for HYP071

Mean continental		Oceanic		Thick continental	
<i>z</i> (km)	<i>v</i> (km s ⁻¹)	<i>z</i> (km)	<i>v</i> (km s ⁻¹)	<i>z</i> (km)	<i>v</i> (km s ⁻¹)
0	6.10	0	6.10	0	6.10
15	6.7	15	7.9	15	6.2
30	8.0	30	8.0	45	8.0
100	8.2	100	8.2	100	8.2
200	8.3	200	8.3	200	8.3
300	8.58	300	8.58	300	8.58

z is the depth, *v* the velocity.

One major problem remains: the depth of the foci. Two deep earthquakes (*h* = 640 km) which do not fit easily in the active tectonics schemes, occurred beneath Granada, in southern Spain (29 March 1954 and 30 January 1973). Moreover, Munuera⁶ listed some intermediate depth earthquakes, but as the seismic network in Morocco and the location procedure were less developed, his locations are not very well established.

Arrival times, of both P and S waves, given in bulletins for stations in Spain, Algeria and Portugal are used here with our data in Morocco for the period 1972–78. The velocity structure used in the HYP071 routine (Table 1) represents an average continental structure as is the case in the Moroccan Meseta⁷. From more than 1,000 earthquakes of magnitude >3 detected in this area, we selected 401 events recorded in 5–25 stations (depending on the magnitude of the earthquake), with a residual smaller than 2 s. Among these we selected 247 earthquakes for which the standard deviation in location or depth, determined by HYP071, was <20 km (Fig. 1).

However, HYP071 error determination is a statistical evaluation of the error based on the dispersion of arrival times but does not take into account systematic bias due to the use of a wrong velocity structure or *V_P/V_S* ratio. We carried out a series of tests, on a few sample events in different regions (Alboran Sea, Gulf of Cadiz, High Atlas, Betic Cordillera, Rif) located with 5–15 arrival times. Starting with a synthetic zero-residual solution, we re-ran the location program, altering one parameter each time. The difference from the initial result gives us an estimate of the error associated with each parameter.

The first test examined the effects of errors in readings. Assuming that misreadings are <1 s for P waves and <2 s for S waves, random noise was added to synthetic travel times. The variations in computed location and depth were respectively <10 km and <20 km in all cases.

The effect of the *V_P/V_S* ratio was also investigated. Again synthetic travel times were used and locations recomputed using *V_P/V_S* ranging from 1.70 to 1.80. The resulting variations in locations and depth were similar to those obtained in the previous test.

The influence of the velocity structure was checked by calculating locations using the two extreme models given Table 1. These models correspond to a thin crust overlying an anomalously low upper mantle as in the Alboran Sea^{5,8} and to a thick crust as in the Betic Cordillera¹⁰ or the Rif zone⁵. Between these models the epicentres differ by 2 km and the depth by 13 km. HYP071 does not take into account lateral variations of velocity but it is unlikely that locations will be much altered by such variations.

One further question is related to the use of a flat layer velocity model, which could not be appropriate over an area ~1,000 km wide, because of the sphericity of the Earth. Chatelain *et al.*⁹ have shown that the resulting error for a 300-km wide network is less than 1 km. We assume the error to be of the same order in our case.

Figure 2a shows an east–west cross-section of the Alboran Sea. Most of the earthquakes are deeper than 50 km. One of them (7 August 1975) was recorded in 198 stations; its depth computed by the ISC is 94 ± 2 km. Observations of pP and sP waves confirm this result⁵. Our determination using 13 stations is 105 ± 4 km. The seismic zone lies between the peridotite

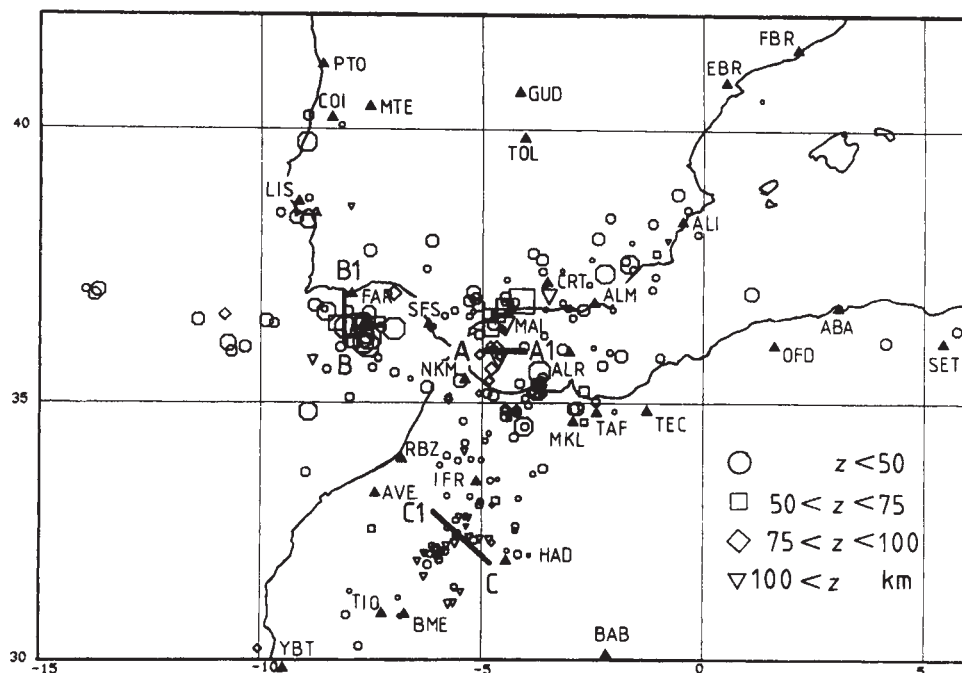


Fig. 1 Seismicity map of the Gibraltar area for the period 1972-78. Errors in location and depth (given by HYP071) are smaller than 20 km. Different symbols are functions of the depth of the focus; the size of the symbol is function of the number of stations which recorded the earthquake. Δ , Stations. The location of the cross-sections are shown.

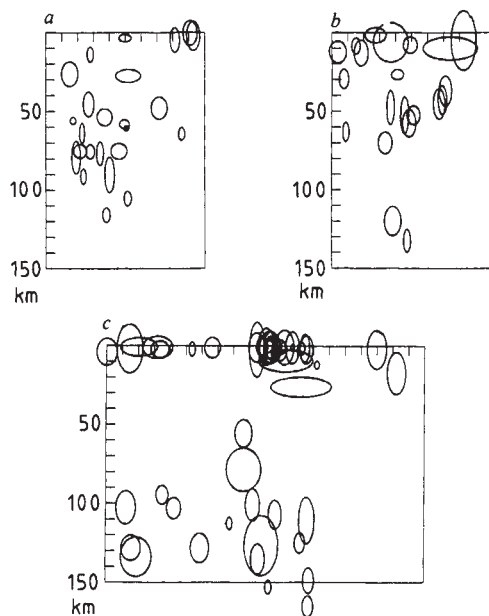


Fig. 2 Cross-sections for the three zones with intermediate earthquakes. *a*, The Western Alboran Sea; *b*, the Gulf of Cadiz; *c*, the High Atlas. The ellipses represent the error in location and depth given by HYP071.

massifs of Ronda (Spain) and Beni Bouchera (Morocco). A refraction profile in this area shows an anomalously high upper mantle velocity⁵ and that relative residuals at nearby stations are large¹¹, corroborating the existence of an anomaly in the velocity structure of this area.

Figure 2*b* shows a north-south cross-section of the Gulf of Cadiz. Most earthquakes there are shallower than 70 km, but three are deeper than 100 km. Commenting on the seismic intensity map of the 15 March 1964 earthquake, Udias and Lopez-Arroyo¹² suggest that its depth should be greater than the 27 km value inferred by the ISC.

In the High Atlas region we recorded 17 earthquakes deeper than 30 km in 7 yr. Because of their small magnitude (<3.5) these earthquakes are usually clearly recorded by only five

Moroccan stations. Thus it is important to have some confidence in the S arrival times. Figure 3 shows typical seismograms for an event at 125 km depth. Because of the significance of intermediate earthquakes in this region, the tests described earlier were applied individually to each event. Figure 2*c* shows a cross-section striking NW-SE across the High Atlas. Of the 17 events, 14 are deeper than 100 km, and there is a gap in activity between 30 and 55 km. The only three events in the 20-90 km range are on the north-east border of the seismic zone.

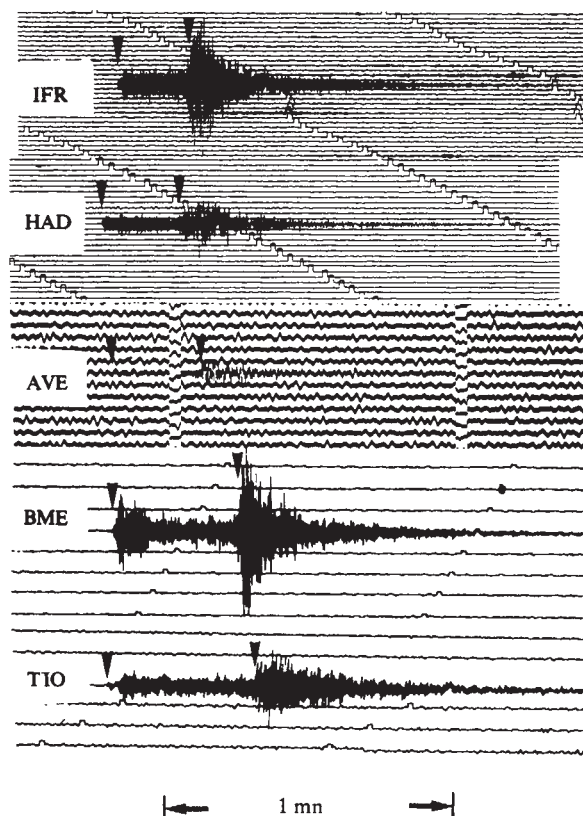


Fig. 3 Seismograms of the 1 August 1977 earthquake located at 125 km depth beneath the High Atlas. Notice the quality of P and S arrival times. Stations names refer to Fig. 1.

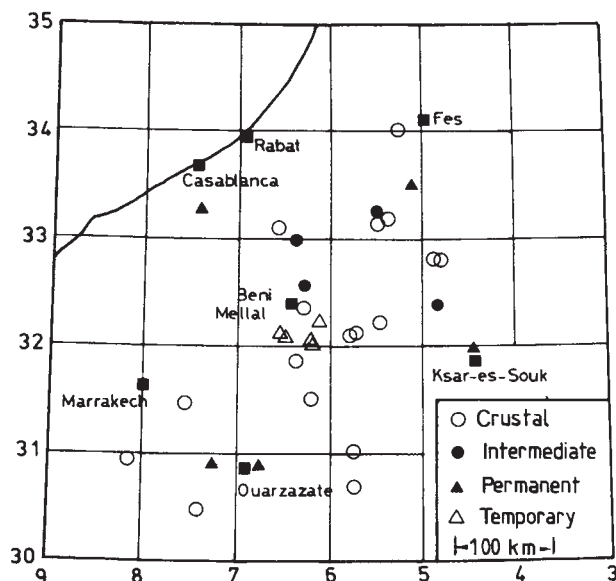


Fig. 4 Epicentres of a 15-day microearthquake survey in the High Atlas. Only events with a difference in arrival of P and S > 5 s and error in depth < 25 km are shown. $\circ$, Earthquakes; $\triangle$, seismological stations.

To check the intermediate seismicity beneath the High Atlas, we carried out a microearthquakes survey for 15 days in 1978¹³. The network of six stations had an aperture of 30 km. During this period, 99 events were recorded in at least four stations; 43 of them have a travel time difference between P and S of < 5 s and are shallower than 20 km. Among the 56 others, 51 are shallower than 20 km, one is 27 km deep and the last four are deeper than 110 km (Fig. 4). Their magnitude ranges between 1.7 and 2.5 and the closest station is at a mean distance of 60 km. The tests described above applied to these four events show a variation of 25 km in depth and 30 km in epicentre.

Two of the three zones where there are reliable intermediate earthquakes are close to the oceanic boundary between Africa and Europa and could be related to subduction (Gulf of Cadiz and Alboran Sea). On the other hand, the High Atlas whose basement and surroundings are continental is considered a typical intracontinental chain¹⁴. The existence of basaltic intrusions of Jurassic age¹⁵ is probably related to an extension episode into an entirely continental surrounding^{14,16}.

Other zones, with intermediate depth earthquakes, are known in the Alpine-Himalayan belt such as Romania, Hindu-Kush, Tibet and Burma. In the Hindu-Kush and Romania a gap in seismicity is observed between 20 and 70 km and the activity decreases at 150 km and disappears at 200 km. This is similar to the results we observe in the High Atlas and differs from that in our two oceanic areas. Although we cannot exclude the possibility that deep seismicity represents the remains of subducting oceanic crust in all of these regions, it is open to question. In the High Atlas particularly, the possibility of a subduction zone in the past 20 Myr seems impossible.

We thank D. Bensari and H. Msougar for help in Morocco, J. Bonnin and M. Mattauer for valuable comments; and especially P. Molnar for his encouragements and suggestions. This research was supported by INAG, contracts ATP 32-01 and 32-23.

Received 30 March; accepted 2 June 1981.

- McKenzie, D. P. *Geophys. J. R. astr. Soc.* **30**, 109 (1972).
- Udias, A. & Lopez-Arroyo, A. *Tectonophysics* **31**, 259 (1976).
- Udias, A. *Rock Mechan. Suppl.* **9**, 75 (1980).
- Hatzfeld, D. *Ann. Géophys.* **32**, 71 (1976).
- Hatzfeld, D. thesis, Univ. Grenoble (1978).
- Munuera, J. M. *El Mapa de Zonas Sísmicas Generalizadas de la Península Ibérica* (Instituto Geográfico y Catastral, Madrid, 1969).
- Hatzfeld, D. & Bensari, D. *Bull. Soc. géol. Fr.* **7**, 749 (1977).

- Hatzfeld, D. C. r. hebdomadaire Séanc. Acad. Sci., Paris **283**, 1021 (1976).
- Chatelain, J. L., Roeckers, S. W., Hatzfeld, D. & Molnar, P. *J. geophys. Res.* **85**, 1365 (1980).
- Banda, E. & Ansorge, J. *Geophys. J. R. astr. Soc.* **63**, 515 (1980).
- Goula, X., DEA, Univ. Grenoble (1976).
- Udias, A. & Lopez-Arroyo, A. *Tectonophysics* **9**, 323 (1970).
- Frogneaux, M. thesis, Univ. Grenoble (1980).
- Mattauer, M., Tapponnier, P. & Proust, F. *Bull. Soc. géol. Fr.* **7**, 521 (1977).
- Hailwood, E. A. & Mitchell, J. G. *Geophys. J. astr. Soc.* **24**, 351 (1971).
- Michard, A. *Notes et Mémoires du Service Géologique, Rabat*, 252 (1976).

Palaeomagnetism of the Ibero-Armorican arc and the Hercynian orogeny in Western Europe

H. Perroud & N. Bonhommet

Laboratoire de Géophysique Interne—Centre Armorican d'Etude Structurale des Socles, (CNRS), Université de Rennes I, Campus de Beaulieu, 35042 Rennes Cédex, France

Van der Voo¹ has pointed out that further advances in the understanding of the Hercynian orogeny require palaeomagnetic data from well studied orogenic areas. We report here palaeomagnetic results of Palaeozoic rocks taken in Brittany, Portugal and Spain in the Ibero-Armorican arc (IAA), a large geological feature considered to be a key area for the Hercynian orogeny in Western Europe. They show evidence that the IAA formed in two stages, extending results obtained in Spain by Ries *et al.*². Giving precise constraints for the former shape of the arc we can use pre-Hercynian palaeomagnetic directions found in Spain and Portugal to follow its evolution during Palaeozoic times. This megastructure linking Brittany to Iberia, squeezed between Gondwana and the northern continents during the Carboniferous, did break away from the former block during Ordovician times.

Geological^{3,4} and geophysical⁵ evidence suggests that the IAA is formed of large juxtaposed curved units lying on the two sides of the Bay of Biscay. Stratigraphic studies of Palaeozoic formations from the arc^{6,7}, performed at the Rennes CNRS centre, suggest that the shape of this structure could have been quite different before the Hercynian deformation. In particular, the very strong correlations between the stratigraphy and palaeontology of the Crozon area (Brittany) and the Buçaco syncline (Portugal) for Ordovician rocks imply that these two regions belonged to the same sedimentary basin in the Lower Palaeozoic, and thus palaeogeographic reconstructions⁷ tend to place these two regions next to each other. Moreover, the question of whether the arc is of primary or secondary origin is crucial for several of the models proposed for the Hercynian orogeny. A primary origin implies that the arc always had the same curvature whereas a secondary origin implies oroclinal bending following Carey's model⁸. Ries *et al.*² recently showed that the Iberian arc, the inner part of the IAA, had a predominant secondary origin. The absence of definitive evidence that the two structures did evolve simultaneously led us to extend the sampling to the whole IAA and to different stages in the Palaeozoic.

Samples were from: (1) Brittany: Upper Ordovician dolerites of Crozon (14 sites, 100 samples); (2) Spain: Ordovician volcanics of Cabo de Penas (2 sites, 31 samples), Upper Devonian red sandstones of Candás (3 sites, 32 samples), Lower Carboniferous red grits, Alba formation from Candás (3 sites, 23 samples) and San Emiliano (6 sites, 42 samples); (3) Portugal: Cambro-Ordovician redbeds (7 sites, 60 samples) and Upper Ordovician dolerites (4 sites, 34 samples) from Buçaco (Fig. 1). Natural remanent magnetization (NRM) was measured for these samples and for most of them a viscosity test⁹ was also performed; a.c. field, thermal and chemical demagnetizations were carried out (all samples being treated by at least one of these methods, mostly by two). In addition, rock magnetic experiments such as coercivity spectrum analysis¹⁰ of pairs of

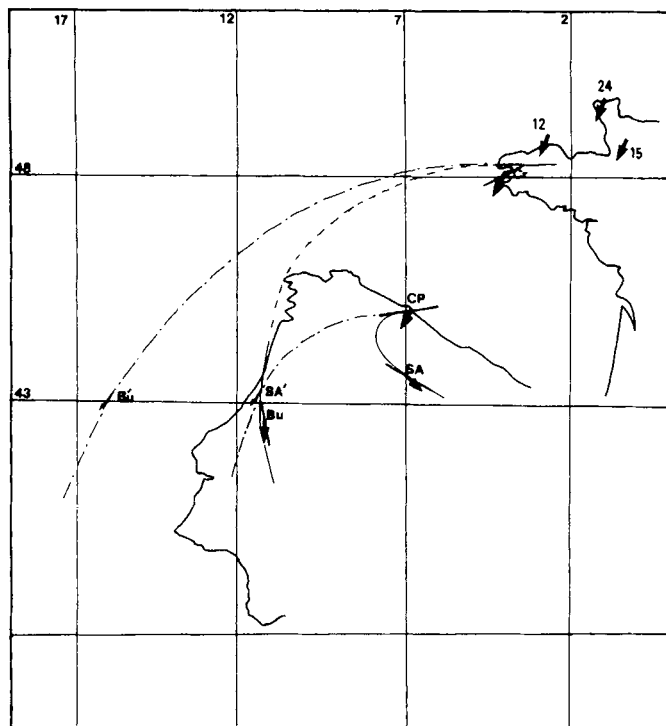


Fig. 1 Palaeomagnetic Carboniferous directions (thick arrows) found on the Ibero-Armorican arc with sampling locations (CR, Crozon; CP, Cabo de Penas and Candas; SA, San Emiliano; BU, Buçaco) after closing of the Bay of Biscay (fitted by Lefort²⁴). Values for additional data in Brittany correspond to the references given in the text (see also Table 1). The ante-tectonic shape of the arc deduced from the data is also given with the corrected Carboniferous directions (small arrows).

leached and unleached samples, heating to Curie temperature or cooling to the liquid air temperature of an acquired induced remanent magnetization (IRM), were carried out to support the interpretation of the multivectorial magnetization with physical arguments on the magnetic carriers. The results obtained from

these studies¹¹, which will be published in detail elsewhere, can be summarized as follows.

The magnetic behaviour of the Upper Ordovician dolerites from Buçaco and Crozon is identical. Intensities are low, with a maximum frequency between 10^{-2} and 10^{-3} A m⁻¹ and viscosities are high (>20% for 50% of the samples); the IRM acquisition curves are identical; they display the same magnetic transition at -160 °C, suggesting the presence of magnetite, and current work at the Rennes CNRS centre gives K-Ar ages which range for both regions from 190 to 300 Myr. In accordance with the radio-dating results, the remanent magnetization is interpreted as being secondary with an acquisition period between Carboniferous and Triassic or Jurassic time (Table 1), thus confirming the geological relationships between Crozon and Buçaco.

Buçaco redbeds exhibit magnetization directions interpreted as pre-tectonic and remagnetization directions as being from the Upper Cretaceous and Quaternary¹¹. A characteristic of the pre-tectonic directions is a constant declination of 150° and variable inclination. The same behaviour was found by Duff¹² in an extensive palaeomagnetic survey of redbeds in Brittany with ages extending from Lower Cambrian to Lower Ordovician. From this work two sites at Crozon are well suited for comparison with Buçaco as both areas are thought to be part of the same sedimentary basin⁷, and thus allow comparison with data from contemporaneous redbeds. Components B3 and C found by Duff¹² on 10 specimens of these two sites using acid leaching, thermal and a.f. cleaning, gave a mean direction of 220° after tectonic correction (sites CP, CR in Tables 2 and 3 of ref. 12). The difference in declination between the two areas suggests a partial secondary origin of the IAA if we correct for the rotation of Spain during the Cretaceous and look at the variation of strike of the arc. Again, the palaeomagnetic characteristics of contemporaneous samples from Brittany and Portugal are identical.

At Cabo de Penas, Ordovician volcanics show two well defined pre-tectonic directions of magnetization (Table 1). One (component A) has a very steep inclination, coercivities over 70 mT and unblocking temperatures over 500 °C, and has been interpreted as Ordovician; the other (component B) has a shallow inclination and is comparable with the direction obtained east of Cabo de Penas for the Upper Devonian and Lower Carboniferous redbeds of Candas (component A); the

Table 1 Palaeomagnetic data from the Ibero-Armorican arc

Site (reference on Fig. 1)	Rock unit	Component of magnetization	N	D	I	k	α_{95}	$\bar{S}$	Age of magnetization	Ref.
South limb										
SA	San Emiliano redgirts	—	30	102	13	53	2	270°	Lower Carboniferous	14
BU	Buçaco dolerites	—	20	139	-43	58	4	145	Jurassic ?	This work
	Buçaco redbeds	A	7	151	80	83	6	145	Ordovician	This work
		B	7	143	61	41	9	145	Between Ordovician and	
		C	6	144	36	45	9	145	Lower Carboniferous ?	
		D	3	153	5	212	6	145	Lower Carboniferous	This work
North limb										
CP	Cabo de Penas volcanics	A	13	208	78	50	6	230	Ordovician	This work
		B	78	178	19	39	8	230	Lower Carboniferous	This work
	Candas redbeds	A	8	185	23	27	10	230	Middle-Upper Carboniferous	This work
CR	Crozon dolerites	A	19	208	-15	215	2	240	Triassic	This work
		B	38	205	29	33	4	240	Middle Carboniferous	This work
		C	12	350	-11	300	3	240	Upper Carboniferous	This work
	Armorican redbeds	B3	10	216	41	35	10		Devonian to early Carboniferous	12
		C	16	228	63	70	7		Silurian to early Carboniferous	12
15	Montmartin redbeds	—	186	208	16	20	7	270	Late Devonian-early Carboniferous	15
24	Flamanville granite	—	8	203	13		14		Carboniferous	22
12	Ploumanach granite (300 Myr)	—	31	200	9	121	7		Middle Carboniferous	12
	Jersey dolerites	A	—	199	16	71	9		Middle-late Carboniferous	16
		B	—	204	57	59	8		Siluro-Devonian	16
	Jersey volcanics (533 Myr)		66	292	78	8	19		Cambrian	16

Palaeomagnetic component directions found on the Ibero-Armorican arc. N, Number of samples used; D, I, declination, inclination; k, α_{95} Fisher statistical parameters²³; $\bar{S}$, mean structural direction. Statistics may refer to samples, specimen, directions or sites: see in each case original work given in the reference; only the total number of samples used is given here; all data were cleaned using a.f., acid and thermal cleaning as done by Duff^{12,16}, N.B. *et al.*¹⁴ and Jones *et al.*¹⁵ and a.c. field up to 0.2 T (2,000 oe) as in ref. 22.

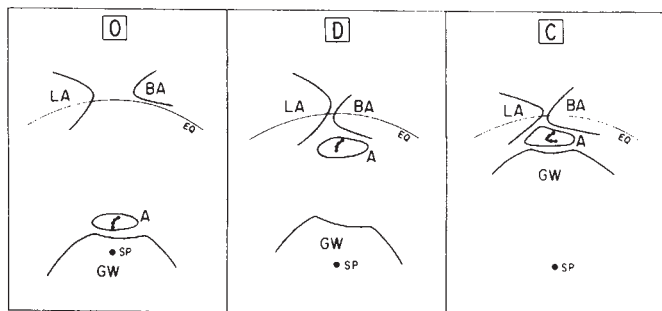


Fig. 2 Scheme of the evolution of the main plates involved in the Hercynian orogeny in Palaeozoic times (O, Ordovician; D, Devonian; C, Carboniferous) with mention of the Ibero-Armorican arc. LA, Laurentia; BA, Baltica; GW, Gondwana; A, Armorica.

eight sites (volcanics and sedimentary) were used to give a mean direction which is pre-tectonic as shown by a fold test. We can do better than the classical tilt correction on this example to calculate the mean direction: if we look at the angle between the strike direction and the magnetization on each site, this angle seems to be very constant ($52.5^\circ \pm 5^\circ$) regardless of the site and even though strike directions vary on this example between 210° and 244° . This suggests that deformation occurred in two stages, the latter one with a vertical axis correlated with the tectonic closure of the arc¹³, and that the magnetization was acquired before deformation; taking this result into account and using a regional mean direction for strike of 230° , the mean palaeomagnetic direction of the eight sites in Table 1 has been obtained after correcting each site for strike variation. The six sites of redbeds near Candas also show a soft component obtained on a few samples but scattered among all the sites (Candas redbeds component B, Table 1); the post-tectonic origin of this component is clearly indicated by a negative fold test¹¹. The corresponding inclination is very close to that obtained at Crozon on the dolerites (Crozon dolerites component B, Table 1); this last component has also been interpreted as post-tectonic but with a magnetization predating the closure of the arc, as shown by using a 'strike direction' correction on the 14 sites of Crozon. Magnetization was therefore acquired at Crozon between two main phases of deformation of the IAA.

The direction given in Table 1 for San Emiliano was obtained by N.B. *et al.*¹⁴ using an unfolding model which took into account the two phases of deformation in the area; the effect of the model was to increase the precision parameter k (from 14 to 64) in comparison with the usual tilt correction. A pre-tectonic direction was found (Table 1) and a Lower Carboniferous age for the magnetization inferred from the age of the sediments.

Ries *et al.*² sampled the same Alba formation at Carranques near Candas and at San Emiliano; however, their result cannot be used for comparison because only three samples at Carranques were thermally cleaned and the authors report a direction $D = 211^\circ$, $I = 48^\circ$ with $k = 2$ showing large scatter after thermal cleaning; this is a general behaviour of all results obtained after cleaning as reported in that study. At San Emiliano four sites have been studied for the NRM but none appears among the thermal cleaned results given in Table 2 of their study.

Carboniferous data from Brittany obtained by Jones *et al.*¹⁵ on the syncline of Montmartin and by Duff¹² from the Ploumanach granite, Jersey dolerites and volcanics, together with our own results at Crozon and Candas (CR, CP, Fig. 1) reveal (Table 1) that the corresponding directions are not significantly different for the northern part of the IAA before, (Candas A) and after (Candas B, Crozon B) deformation: most of the rotation seems to have occurred in the southern part of the arc. A new shape of the IAA can then be inferred taking the simplest compatible solution with the palaeomagnetic data. This

indicates that 80° of the curvature of the arc are secondary and that 70° are primary or pre-Carboniferous.

Ries *et al.* used results based on NRM directions from 29 sites along the Iberian arc to demonstrate convincingly the partially secondary origin of this structure; however, because these authors did not consistently use cleaned results to look for multicomponent magnetization, because they infer a Variscan age for all magnetizations, their data cannot be used for comparison as they give no directions associated with the age of magnetization for the different rock units studied. However, their result, a bending of 110° , is not very different from the 80° reported here.

From the new position of the sampling areas and the palaeomagnetic vectors calculated for the new shape of the arc (that is, after elimination of the Hercynian deformation), we can determine virtual palaeomagnetic poles for our Ordovician sites. The Buçaco (25°N , 335°E) and Cabo de Penas (30°N , 330°E) poles coincide in the new reconstruction, indicating that the shape of the arc we deduced from the Carboniferous data is also compatible with our Ordovician results. We therefore propose that 70° of the arc are effectively primary (presedimentary), inherited from Precambrian times. The mean Carboniferous declination of the north limb of the IAA differs by a few degrees (10° – 15°) from that of stable Europe as shown by Jones *et al.*¹⁵ for Montmartin syncline; it is not clear, however, whether the bending took place entirely in the southern part of the arc with a further rotation of 15° of the whole Ibero-Armorican block or whether it was the closure of the arc that caused the 15° difference with Europe. However, this does not change the former shape of the IAA found from the data. The Crozon B and Candas B components, both post-tectonic but ante-closure of the arc, in agreement with results of Jones *et al.* at Montmartin, show that at that time the IAA with its primary shape was close to Europe, with a large ocean separating it from Gondwana, a result shown for Brittany by Jones *et al.*¹⁵ (Fig. 2).

The Ordovician mean palaeomagnetic pole obtained (27°N , 332°E) falls between Cambrian and Siluro-Devonian poles found by Duff¹⁶ for the Armorican massif and this agrees with the Ordovician age assumed for the magnetization. Interestingly, this pole is quite different from conventional Ordovician poles for the British Isles¹⁷ but could more easily be correlated with north Wales data, which have until now been considered as anomalous¹⁸. Recent results by Hagstrum *et al.*¹⁹ showing that the apparent polar wander curve of the Armorican massif is very similar to the Eo-Cambrian–Cambrian curve for Gondwanaland prompted us to compare the Ordovician palaeolatitude for the IAA (70°) with the African Ordovician pole²⁰; this shows that the IAA should have been very near to Gondwanaland in the early Palaeozoic. These results show, moreover, that the Armorican massif and Iberia Meseta were probably a single unit during the entire Palaeozoic, thus justifying the name of 'Armorica plate' given by Van der Voo²¹ to these regions of Western Europe; the large drift of this plate since the Ordovician shows that collision and subduction tectonic processes did exist during the Hercynian orogeny. Finally, taking into account all these results and using available data given by Van der Voo¹, we propose a scheme of the main plates involved in the Hercynian orogeny with particular reference to the Armorica plate bearing the IAA (Fig. 2).

The research was supported by a grant from INAG through 'ATP Géodynamique' national programme. We thank Dr Van der Voo for suggestions and comments.

Received 2 February; accepted 2 June 1981.

1. Van der Voo, R. *et al.* *Int. geol. Congr. Colloq.* C6, 204–212 (1980).
2. Ries, A. C. *et al.* *Earth planet. Sci. Lett.* **50**, 301–310 (1980).
3. Cogné, J. *Histoire Structurale du Golfe de Gascogne* (IFP) 11.1, 11–23 (1971).
4. Matte, P. *Nova Acta Leopoldina* **224**, 239–262 (1974).
5. Lefort, J. P. *Geology* **7**, 384–388 (1979).
6. Henry, J. L. *et al.* *Commun. Serv. Géol. Port.* **57**, 303–345 (1974).
7. Paris, F. & Robardet, M. *Bull. Soc. Géol. Fr.* **7**, 1121–1126 (1977).
8. Carey, S. W. *The Tectonic Approach to Continental Drift* (ed. Carey, S. A.) (University of Tasmania, 1958).
9. Thellier, E. & Thellier, O. *Ann. Geophys.* **15**, 285–376 (1959).
10. Dunlop, D. J. *Geophys. J. R. astr. Soc.* **27**, 37–56 (1972).

11. Perroud, H. thesis, Univ. Rennes (1980).
12. Duff, B. A. *Geophys. J. R. astr. Soc.* **59**, 345–365 (1979).
13. Julivert, M. *Histoire Structurale du Golfe de Gascogne (IFP) 12-1, 12-28* (1971).
14. Bonhomme, N. et al. *J. geophys. Res.* **86**, 1873–1887 (1981).
15. Jones, M. et al. *Geophys. J. R. astr. Soc.* **58**, 287–308 (1979).
16. Duff, B. A. *Geophys. J. R. astr. Soc.* **60**, 355–375 (1980).
17. Piper, J. *Earth planet. Sci. Lett.* **44**, 176–192 (1979).
18. Thomas, C. & Briden, J. C. *Nature* **259**, 380–382 (1976).
19. Hagstrum, J. T. et al. *Geophys. J. R. astr. Soc.* **61**, 489–517 (1980).
20. Morel, P. & Irving, E. J. *Geol.* **86**, 535–561 (1978).
21. Van der Voo, R. *EOS Trans.* **60**, 241 (1979).
22. Van der Voo, R. & Klootwijk, C. T. *Geologie Mijn.* **51**, 609–617 (1972).
23. Fisher, R. A. *Proc. R. Soc. A* **217**, 295–305 (1953).
24. Lefort, J. P. *Mar. Geol.* **37**, 355–369 (1980).

Dietary carbon sources of mussels and tubeworms from Galapagos hydrothermal vents determined from tissue ^{14}C activity

P. M. Williams*, K. L. Smith*, E. M. Druffel† & T. W. Linick†

* Scripps Institution of Oceanography and

† Department of Chemistry, University of California, San Diego, La Jolla, California 92093, USA

The large quantities of reduced carbon that are required to support the filter-feeding mytilid mussels (*Mytilus* sp.), vesicomyid clams (*Calymene* sp.) and various other animals in the Galapagos hydrothermal vent systems are thought to be derived from either the *in situ* synthesis of particulate organic matter by chemoautotrophic, sulphide-oxidizing bacteria^{1,2} or by the advection of sedimentary organic carbon into the vent environment from surrounding areas^{3,4}. In contrast, the dense populations of vestimentiferan tubeworms (*Riftia pachyptila*), which lack mouth organs and digestive tracts, apparently utilize organic carbon synthesized by symbiotic chemoautotrophs⁵. We present evidence here, based on ^{14}C activities and $^{13}\text{C}/^{12}\text{C}$ ratios, that the principal source of dietary carbon for mussels and tubeworms is derived from the dissolved inorganic carbon (DIOC) in the vent effluent waters.

There are two postulated sources of particulate organic carbon (POC) available as dietary carbon to filter-feeding organisms. The first is sedimentary POC derived from DIOC fixed photosynthetically at the ocean's surface which subsequently reaches the ocean floor. Its radiocarbon activity is assumed to be the same as that of the DIOC in surface water ($+20 \pm 20$, Table 1). The $\delta^{13}\text{C}$ of this POC (-22 , Table 1) is a mean value derived from $\delta^{13}\text{C}$ measurements of POC^{6,7} collected below 500 m. The second source is POC synthesized chemoautotrophically from DIOC in the vent waters. This DIOC originates from both magmatic activity⁸ ($\Delta^{14}\text{C} = -1,000$ or 'dead' carbon) and the ambient bottom water ($\Delta^{14}\text{C} = -233$, Table 1) which mixes with the high-temperature hydrothermal fluid. As the organic ^{14}C activities of the organisms grown in the vent systems reflect the $\Delta^{14}\text{C}$ of these three carbon sources and the $^{13}\text{C}/^{12}\text{C}$ ratios reflect the carbon isotope fractionation occurring during synthesis of organic tissues in these organisms, it is possible to calculate the relative contribution of each carbon source using the following equations:

$$\Delta^{14}\text{C}_{\text{mussel}} = x(-1,000) + y(-233) + z(20) \quad (1)$$

$$\delta^{13}\text{C}_{\text{mussel}} = (x + y)\delta^{13}\text{C}_{\text{vent POC}} + z(-22) \quad (2)$$

where x , y and z are the relative amounts of magmatic DIOC, ambient DIOC and sedimentary POC, respectively. The $\delta^{13}\text{C}$ of POC in the vent water is unknown. Both of the above equations have to be satisfied to determine the relative amounts of the three possible dietary carbon sources utilized by the filter-feeding organisms.

The quantitative contribution of sedimenting POC to the food supply for filter-feeding vent organisms is questionable considering the flux of POC to the ocean floor and the *in situ* respiration rates of mussels. The total, mean flux of POC to the sea floor at 2,670 m $0^{\circ}35.75' \text{N}$, $86^{\circ}05.66' \text{W}$ measured 20 and 100 m above the bottom on 4 July 1976 was $2.0 \text{ g C m}^{-2} \text{ yr}^{-1}$ (ref. 9). *In situ* measurements of individual mussel respiration rates by K. Smith (unpublished results) range from 0.30 to $0.81 \text{ ml O}_2 \text{ h}^{-1}$ in dense mussel beds. Assuming a respiratory quotient of 0.85, then these rates are equivalent to an organic carbon requirement of $1.2\text{--}3.3 \text{ g C yr}^{-1}$ (mean = 2.3 g C yr^{-1}). Thus, each mussel would essentially require the total annual POC flux per square metre for maintenance with no growth.

The ^{14}C activity of the mussel tissue could result from a mixture of sedimentary POC and POC of low ^{14}C activity derived from DIOC in the vent waters. The maximum amount of magmatic DIOC in the Galapagos vent waters ($T = 17^{\circ}\text{C}$) has been estimated to be $375 \mu\text{mol kg}^{-1}$, a 13% increase above the ambient DIOC concentration of $\sim 2,400 \mu\text{mol kg}^{-1}$ (refs 10, 11). If 13% of the vent water DIOC is magmatic ($\Delta^{14}\text{C} = -1,000$) and the remaining 87% is composed of ambient DIOC ($\Delta^{14}\text{C} = -233$) plus sedimentary POC derived from the DIOC in surface water ($\Delta^{14}\text{C} = +20$), then the maximum amount of sedimentary POC that could be incorporated into mussel tissue ($\Delta^{14}\text{C} = -267$) is 26% (equation 1). Estimates of mussel density in the 'Musselbed' vent area (W. Smithy, personal communication) are 20 ± 9 specimens per 0.25 m^2 in the immediate vicinity of the vent plume. If the mussels are utilizing the calculated maximum of 26% of sedimentary POC for growth (taken here as equivalent to maintenance), then each vent musselbed would use $\sim 48 \text{ g C m}^{-2} \text{ yr}^{-1}$ ($0.26 \times 2.3 \text{ g C yr}^{-1}$ per mussel $\times 80$ mussels per m^2), or about 35% of the annual primary productivity per square metre in the euphotic zone⁹ for the growth of mussels alone.

There is a significant difference between the $^{13}\text{C}/^{12}\text{C}$ ratios, as well as ^{14}C activities, of mussel tissue collected 1 and 8 m from the vent plume (Table 1). If sedimentary POC ($\delta^{13}\text{C} = -22$) was incorporated into the tissue of the 8-m mussel, the resulting $\delta^{13}\text{C}$ value (-32.3) would be less than that observed (-32.8). This calculation assumes that the ^{14}C activity of the 8-m mussel tissue ($\Delta^{14}\text{C} = -228$) is derived from a mixture of 14% sedimentary POC ($\Delta^{14}\text{C} = +20$) and 86% chemosynthetic POC ($\Delta^{14}\text{C} = -267$) (equation 1).

If the ^{14}C activity of mussel shell carbonate reflects the ^{14}C activity of DIOC in the surrounding seawater, then the 21% decrease in $\Delta^{14}\text{C}$ between the ambient seawater DIOC and the mussel shell carbonate reflects dilution of the ambient DIOC with 3% magmatic DIOC. For the two clam shell carbonates, the 30% decrease in $\Delta^{14}\text{C}$ would represent dilution of the ambient DIOC with 4% magmatic DIOC, taking the ^{14}C activity of the bottom water at 21°N to be the same as at the Galapagos rift zone. A similar calculation gives 4.4% magmatic DIOC incorporated into cellular carbon synthesized by the chemoautotrophic bacteria and used by the mussels 1 m from the plume (provided the contribution from sedimentary POC is negligible).

The identical ^{14}C activities of the vestimentiferan tubeworm tissue and the mussel tissues collected 1 m from the vents strongly suggest that the tubeworms and mussels are utilizing the same DIOC sources, even though $\delta^{13}\text{C}$ in the tubeworm tissue is 23% greater than in the mussel tissue (Table 1). These atypical $^{13}\text{C}/^{12}\text{C}$ ratios found in the vent organisms have been discussed by Rau^{12,13}. If bacterial chemosynthesis is the major process providing dietary carbon for the mussels and tubeworms, it is still unclear why carbon isotope fractionation should be different between chemoautotrophic synthesis occurring in the vent waters and internally (symbiotically) in the tubeworms.

An additional carbon source which may contribute to the bacterial synthesis of POC in the vent waters is methane, energetically a more favourable substrate for the growth of heterotrophic bacteria than DIOC. Methane concentrations up to $3.4 \mu\text{mol kg}^{-1}$ have been measured in the Galapagos vent

Table 1 $\Delta^{14}\text{C}$ and $\delta^{13}\text{C}$ values of Galápagos vent organisms and relevant carbon sources

Sample description	(‰) $\Delta^{14}\text{C}$	(‰) $\delta^{13}\text{C}$	Ref.
Mussel tissue—dive 880 ('Musselbed', 1 m from vent)	-270 ± 6	-33.9	This work
Mussel tissue—dive 895 ('Musselbed', 1 m from vent)	-263 ± 8	-33.8	This work
Mussel tissue—dive 880 ('Musselbed', 8 m from vent)	-228 ± 12	-32.8	This work
Mussel shell—dive 991 ('Musselbed', 1 m from vent)	-254 ± 6	$+2.8$	This work
Tubeworm tissue—dive 993 ('Garden of Eden', in vent plume)	-270 ± 20	-10.9	This work
Clam shell—dive 981 (East Pacific Rise, 21°N)	-263 ± 6	$+3.1$	This work
Mussel tissue ('Clambake I' specimen, 8 replicates)	—	-32.7 to -33.6	12
Tubeworm tissues ('Rose Garden')	—	-10.8 to -11.0	13
Clam shell ('Clambake I')	-263 ± 10	—	10
Total DIOC, 2584 m* (Geosecs Station 337)	-233.4	—	14
Total DIOC, surface sea water	$+20 \pm 20$	-1.7 to -3.5	15
Particulate organic detritus $> 500\text{ m}$	—	-20 to -25 ; mean -22	6, 7

The three mussels for tissue analysis were collected 1 and 8 m from the 'Musselbed' vent plume, *Alvin* dives 880 (21 January 1979) and 895 (20 February 1979), $0^\circ47.89'\text{N}$, $86^\circ09.21'\text{W}$ at 2,493 and 2,480 m. The mussel for shell analysis was collected 1 m from the 'Musselbed' vent plume, *Alvin* dive 991 (8 December 1979) at 2,490 m. The tubeworm was collected in the 'Garden of Eden' vent plume, *Alvin* dive 993 (10 December 1979), $0^\circ47.69'\text{N}$, $86^\circ07.74'\text{W}$ at 2,518 m. The clam for shell analysis was collected from the East Pacific Rise hydrothermal vent system, *Alvin* dive 981 (5 November 1979), $20^\circ50'\text{N}$, $109^\circ06'\text{W}$ at $\sim 2,600\text{ m}$. The mussel tissue and tubeworm specimens were frozen after collection, and the tissue and shell samples rinsed with $^{14}\text{CO}_2$ -free, organic carbon-free distilled water before drying. To determine $\% \Delta^{14}\text{C}$ and $\% \delta^{13}\text{C}$, all samples but the tubeworm tissue were burned in O_2 or acidified with HCl to give CO_2 which was converted to acetylene, via lithium carbide¹⁵ and then counted for 2 days in each of two stainless-steel 1.0-l and 0.4-l gas proportional counters. The worm tissue CO_2 was purified by absorption onto CaO , diluted with background gas and counted twice in a 200-ml copper proportional counter for a total of 12 days. All countings were corrected for isotope fractionation (to a standard value of $\delta^{13}\text{C} = -25.0\%$ PDB-1) and for decay since the time of formation (to AD 1950). The standard was 95% of the net National Bureau of Standards oxalic acid count rate, corrected to a $\delta^{13}\text{C} = -19.0\%$, and the results are reported in terms of $\Delta^{14}\text{C}$ ($\pm 1\sigma$ counting error) which is the per million (‰) deviation from the ^{14}C activity of nineteenth-century wood¹⁶. The $^{13}\text{C}/^{12}\text{C}$ ratios are reported as $\delta^{13}\text{C}$ ($\pm 0.2\%$) relative to the PDB-1 standard.

* The $\Delta^{14}\text{C}$ of the 2,490-m bottom water at the 'Musselbed' was assumed to be identical to the $\Delta^{14}\text{C}$ in 2,584-m water at Geosecs station 337, 24 May, 1974, $04^\circ50'\text{N}$, $124^\circ05'\text{W}$. σ_T at 2,500 m at both locations is 27.76 ± 0.01 (ref. 17).

† $\Delta^{14}\text{C}$ values taken from corals collected from 3–5-m depth at 1.5°S , $90^\circ\text{–}91^\circ\text{W}$, assuming the $\text{Ca}^{14}\text{CO}_3$ reflects the ^{14}C activity of the surface seawater DIOC.

waters (M. D. Lilly, M. A. de Angelis and L. I. Gordon, unpublished results), and it has been calculated that $50\text{ }\mu\text{mol kg}^{-1}$ of abiogenic CH_4 of magmatic origin could be present in the 'end member' 350°C hydrothermal fluid at the East Pacific Rise¹⁸. If methane were the sole carbon source for the synthesis of POC utilized by the mussels, then the mussel tissue would contain no measurable radiocarbon activity. Obviously this is not the case. It is possible that the ^{14}C activity of the mussel tissues (-267) is 13% lower than that of the mussel shell carbonate (-254) due to mussels taking up both 'dead' CH_4 and DIOC, whereas the shell is only using 'dead' DIOC.

On the basis of the ^{14}C activities and $^{13}\text{C}/^{12}\text{C}$ ratios reported here, we conclude that: (1) filter-feeding organisms in the vent system are directly or indirectly incorporating 'dead' carbon of magmatic origin into their tissue; (2) $\sim 25\%$ or less of the dietary carbon available to the mussels is from sedimenting POC fixed photosynthetically at the surface; and (3) mussel tissue is incorporating relatively more 'dead' DIOC than is mussel shell carbonate in specimens collected at the same location near the vent.

This article is contribution no. 21 of the Galápagos Rift Biology Expedition, supported by the US NSF. We thank G. Somero, K. Turekian and R. Lutz for providing the specimens used in this study, and S. Griffin and C. Hutto for preparation of ^{14}C samples. This work was supported by grants from the US Department of Energy (DE-AMO3-76SF00010) and NSF (OCE78-08856, OCE79-17652).

Received 10 November 1980; accepted 22 May 1981.

- Jannasch, H. W. & Wirsén, C. O. *Bioscience* **29**, 592–598 (1979).
- Karl, D. M., Wirsén, C. O. & Jannasch, H. W. *Science* **207**, 1345–1347 (1980).
- Lonsdale, P. *Deep-Sea Res.* **24**, 852–863 (1977).
- Enright, J. T., Newman, W. A., Hessler, R. R. & McGowan, J. A. *Nature* **289**, 219–221 (1981).
- Cavanaugh, C., Gardner, S. L., Jones, M. L., Jannasch, H. W. & Waterbury, J. B. *Science* (in the press).
- Williams, P. M. & Gordon, L. I. *Deep-Sea Res.* **17**, 19–27 (1970).
- Eadie, B. J., Jeffrey, L. M. & Sackett, W. M. *Geochim. cosmochim. Acta* **42**, 1265–1269 (1978).
- Corliss, J. B. *et al. Science* **203**, 1073–1083 (1979).
- Cobler, R. & Dymond, J. *Science* **209**, 801–803 (1980).
- Turekian, K., Cochran, J. K. & Nozaki, Y. *Nature* **280**, 385–387 (1979).
- Craig, H. & Weiss, R. F. *J. geophys. Res.* **75**, 7641–7647 (1970).
- Rau, G. H. & Hedges, J. I. *Science* **203**, 648–649 (1979).
- Rau, G. H. *Science* (in the press).

- Ostlund, H. G., Brescher, R., Oleson, R. & Ferguson, M. J. *Univ. Miami Rosenstiel School Mar. Atmos. Sci. Tritium Lab. Data Rep.* **8**, (1979).
- Druffel, E. M. *Geophys. Res. Lett.* **8**, 59–62 (1981).
- Broecker, W. S. & Olson, E. A. *Radiocarbon* **3**, 176–204 (1961).
- Wooster, W. S. *Prelim. Rep. Step-1 Exped. 15 Sept.–14 Dec. 1960, Part I. Phys. chem. Data*, SIO ref. no. 61-9 (1961).
- Welhan, J. *Trans. Am. geophys. Union (EOS)* **61**, 992 (1980).

A natural biological control of Dutch elm disease

Joan Webber

Department of Botany and Microbiology, University College of Wales, Aberystwyth SY23 3DA, Dyfed, UK

Since its arrival in the late 1960s, the aggressive strain of *Ceratocystis ulmi*, the fungus that causes Dutch elm disease, has destroyed over 20 million elms in Britain and subsequently inflicted similar heavy losses across much of continental Europe^{1,2}. Successful control of the disease has been achieved only locally, using intensive sanitation and fungicide injection programmes^{3,4}. However, it has recently become apparent that disease spread may also be limited naturally. I present here evidence of a biological control of Dutch elm disease which could be exerting an important effect in some parts of Britain. This control process acts by preventing successful breeding of scolytid beetles which are the vectors of *C. ulmi*.

The control phenomenon was first encountered during 1976 and 1977 in several stands of trees in mid and south Wales. Each stand contained between 10 and 100 wych elms (*Ulmus glabra*) and although *C. ulmi* had killed some trees in all these stands, disease spread to healthy trees was surprisingly slow, suggesting that only small numbers of beetle vectors were present. Investigation of beetle breeding material, that is, the inner bark of dying or recently dead trees, revealed only a limited amount of active breeding. Instead such bark commonly exhibited dark zone lines (Fig. 1a), which on isolation yielded the fungus

Phomopsis oblonga. The presence of *Phomopsis* was also clearly associated with the disruption of beetle breeding and disruption usually took one of two forms. Most commonly, when *Phomopsis* and a breeding gallery occupied the same area of bark, beetle larvae were retarded in their development and constructed very elongated, contorted galleries. Less commonly observed was the evasive action of larvae on encountering a discrete *Phomopsis* lesion. In such circumstances larval galleries swerved away from the *Phomopsis*-colonized area into uncolonized tissue.

To compare effectiveness of scolytid breeding in *Phomopsis*-colonized and uncolonized bark, breeding experiments were done using the two British species of beetle vector, *Scolytus scolytus* and *Scolytus multistriatus*. The breeding material was of two types: logs cut from wych elm with recent, extensive bark colonization by *Phomopsis*, and control logs also from wych elm but with no observable *Phomopsis* colonization. Pairs of logs (one of each type) were presented to beetles in either a 'choice' or a 'forced' breeding situation. In the former, a standing pair of logs was enclosed in a fine net bag and beetles introduced. This allowed beetles to choose as breeding material either the *Phomopsis*-colonized log or the control log. In the forced breeding situation each log was enclosed individually in a bag, so that introduced beetles had no choice of breeding material. The

Table 1 Effect of *Phomopsis* colonization on scolytid breeding in elm bark

Breeding situation in log pairs	Maternal galleries per 100 cm ² bark	Total no. of emergent beetle progeny
<i>S. scolytus</i>		
Forced breeding		
Uncolonized control log	1.0	756*
<i>Phomopsis</i> -colonized log	1.3	0*
Choice breeding		
Uncolonized control log	1.8	1,690*
<i>Phomopsis</i> -colonized log	0.7	0*
<i>S. multistriatus</i>		
Forced breeding		
Uncolonized control log	2.5	780
<i>Phomopsis</i> -colonized log	2.3	17*
Choice breeding		
Uncolonized control log	4.3	1,048
<i>Phomopsis</i> -colonized log	1.7	0*

Four pairs of logs were used in all. The mean area of bark on logs presented to *S. scolytus* was 3,286 cm², and on logs presented to *S. multistriatus* was 2,290 cm². The number of maternal galleries per 100 cm² bark is a mean.

* Some living larvae (<20) still remaining in log.

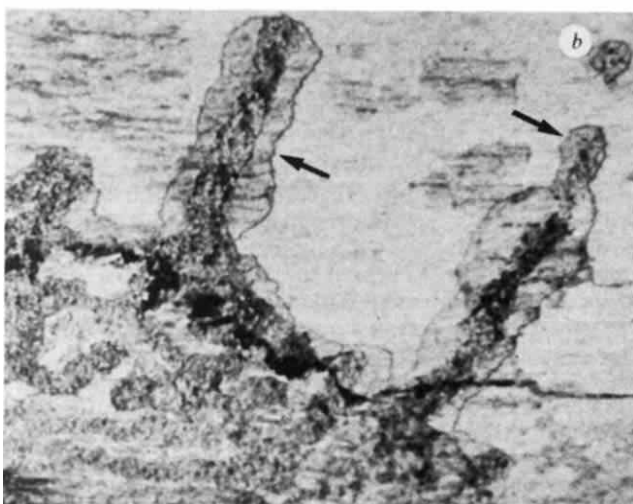
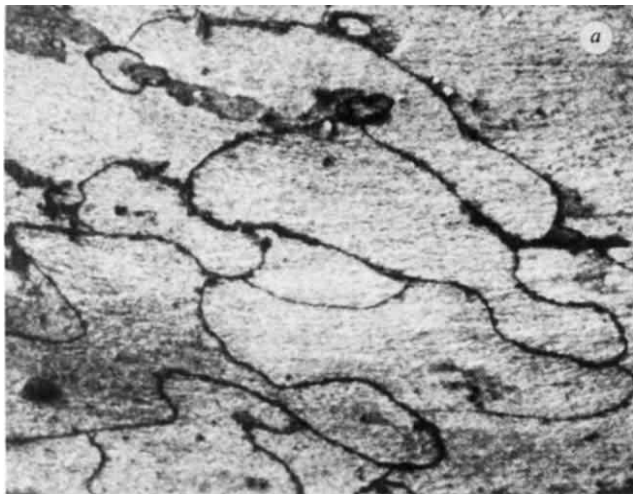


Fig. 1 *a*, Characteristic zone lines produced by *Phomopsis* colonizing the inner bark of diseased elms. These dark lines are thought to result from antagonistic interactions between different forms of the fungus present in the bark⁷. $\times \sim 2$. *b*, Modified zone lines (arrowed) formed around the larval galleries of *Scolytus*. Such modified zone lines are probably formed by *Phomopsis* in response to changes in the moisture content of the bark as a result of gallery excavation. $\times \sim 2.4$.

numbers of beetles introduced in these conditions depended on the total bark area available and the beetle species used. In the case of *S. scolytus*, one beetle pair, consisting of male and female, was allotted to each 100 cm² of bark, whereas one pair per 25 cm² of bark was used for the smaller bark beetle, *S. multistriatus*.

After 6 weeks the beetle breeding cycles were completed and the bark was stripped from the logs and gallery formation compared. It was immediately apparent that the breeding of both *Scolytus* species was strongly influenced by *Phomopsis*. When female beetles attempted to cut maternal galleries in a *Phomopsis*-colonized log many were abandoned at various stages before completion. If galleries were completed and eggs laid, the number of resulting larvae was drastically reduced and few developed into new adult beetles. In contrast, breeding in control logs was successful, and many young adult beetles emerged (Table 1). Figure 2*b* clearly shows the attempts at maternal gallery excavation in bark colonized by *Phomopsis*. Abnormal larval galleries are evident, typified by their elongated, twisted nature when compared with larval galleries in uncolonized bark (see Fig. 2*a*). Galleries affected by *Phomopsis* were also characterized by their frequent association with modified zone lines. Rather than appearing as the irregular ovals shown in Fig. 1*a*, zone lines often enclosed an entire gallery system, outlining each larval mine separately (Fig. 1*b*).

Behaviour of beetles in choice breeding experiments also suggested that *Phomopsis*-colonized bark was usually avoided as breeding material when an alternative was offered. The mean number of maternal galleries per 100 cm² of bark was similar in both *Phomopsis* and control logs in forced conditions, whereas in choice breeding experiments the number of galleries cut in control logs was over twice that attempted in *Phomopsis* logs (Table 1).

Thus *Phomopsis* is clearly able to exert a twofold effect on scolytid beetles by decreasing numbers of viable offspring in brood trees, or by eliminating potential breeding material and therefore intensifying competition for any uncolonized bark. Both factors must invariably have some role in reducing vector populations but we can only speculate on how much this may affect the continuing spread of Dutch elm disease.

One of the major factors which must undoubtedly govern the effectiveness of *Phomopsis* as a control agent is its frequency of occurrence, possibly as a latent inoculum in the dead outer bark of healthy trees. High levels of such inoculum would favour



Fig. 2 *a*, Breeding galleries of *S. multistriatus* produced in the bark of wych elm (*U. glabra*) with no *Phomopsis* colonization. Typically each gallery consists of rows of larval mines (or galleries) radiating outwards from points of egg laying along the length of a central mother gallery. Many of the larval mines terminate in oval pupal chambers. $\times \sim 0.25$. *b*, Sparse and abnormal breeding galleries of *S. multistriatus* in the bark of wych elm colonized by *Phomopsis*. Such galleries commonly show abortive or elongated and contorted larval mines. $\times \sim 0.25$.

rapid colonization of trees if they succumbed to Dutch elm disease. *Phomopsis* has frequently been found in the outer bark of healthy elms, but survey data have shown that it is largely limited to the elms in western and northern parts of Britain (J. F. W. and J. N. Gibbs, in preparation). It is therefore more likely to contribute to natural disease control in these areas. Furthermore, the massive colonization of elm bark required to exert an appreciable effect on beetle breeding has only been observed in wych elm, and it seems that *Phomopsis* is more strongly associated with this species of elm than any other in Britain. The fact that wych elm comprises the main part of elm populations in northern England, Wales and Scotland, also suggests that *Phomopsis* may be effective only in the north and west.

It has long been known that the spread of Dutch elm disease has been slower in northern England and in Scotland⁵. Previously this has been ascribed to various factors, including a smaller and more scattered elm population, a climate less favourable to the vectors and the preponderance of wych elm rather than English elm (*Ulmus procera*), the latter predominating in the south^{3,6}. It now seems that *Phomopsis* in an additional, and in some areas possibly the major factor in slowing the spread of the disease in the north.

I thank Drs J. N. Hedger and J. N. Gibbs for advice and encouragement. This work was supported by a SRC CASE studentship award.

Received 26 May; accepted 6 June 1981.

1. Gibbs, J. N. & Brasier, C. M. *Nature* **241**, 38–383 (1973).
2. Brasier, C. M. *Nature* **281**, 78–80 (1979).
3. Gibbs, J. N. *Arboric. J.* **3**, 110–114 (1978).
4. *Dutch Elm Disease Control* (East Sussex County Council Publ. No. P/131, 1977).
5. Pearce, T. R. *Bull. For. Comm.* **33** (1960).
6. Redfern, D. B. *Scott. For.* **31**, 105–109 (1977).
7. Webber, J. F. thesis, Univ. Wales (1980).

Regenerative and passive membrane properties of isolated horizontal cells from a teleost retina

Daniel Johnston* & Dominic Man-Kit Lam†

Program in Neuroscience, *Department of Neurology and †Cullen Eye Institute, Baylor College of Medicine, Houston, Texas 77030, USA

Neurons in the intact retina are interconnected by a complex network of chemical and electrical synapses. The electrical interactions between many types of retinal neurones make it difficult to determine the active and passive membrane properties of the individual cells. A direct way to examine such properties is to dissociate the retina into single cells^{1–3} and to record intracellularly from identified, isolated neurones in the absence of neural connections. Such an approach has been used to study the electrophysiological properties of several classes of neurones in the vertebrate retina^{3–5}. We have combined the method developed previously for examining the neurochemistry of individual retinal cells^{1,2} with a single-electrode clamp technique^{6,7} to investigate the biophysical properties of isolated horizontal cells from the teleost retina. Our results suggest that the specific membrane capacity of these cells is $\sim 1 \mu\text{F cm}^{-2}$, and that the specific membrane resistivity ranges from $\sim 5,000$ to $20,000 \Omega \text{ cm}^2$. In isolation, these cells can show regenerative voltage responses which are most probably calcium-dependent.

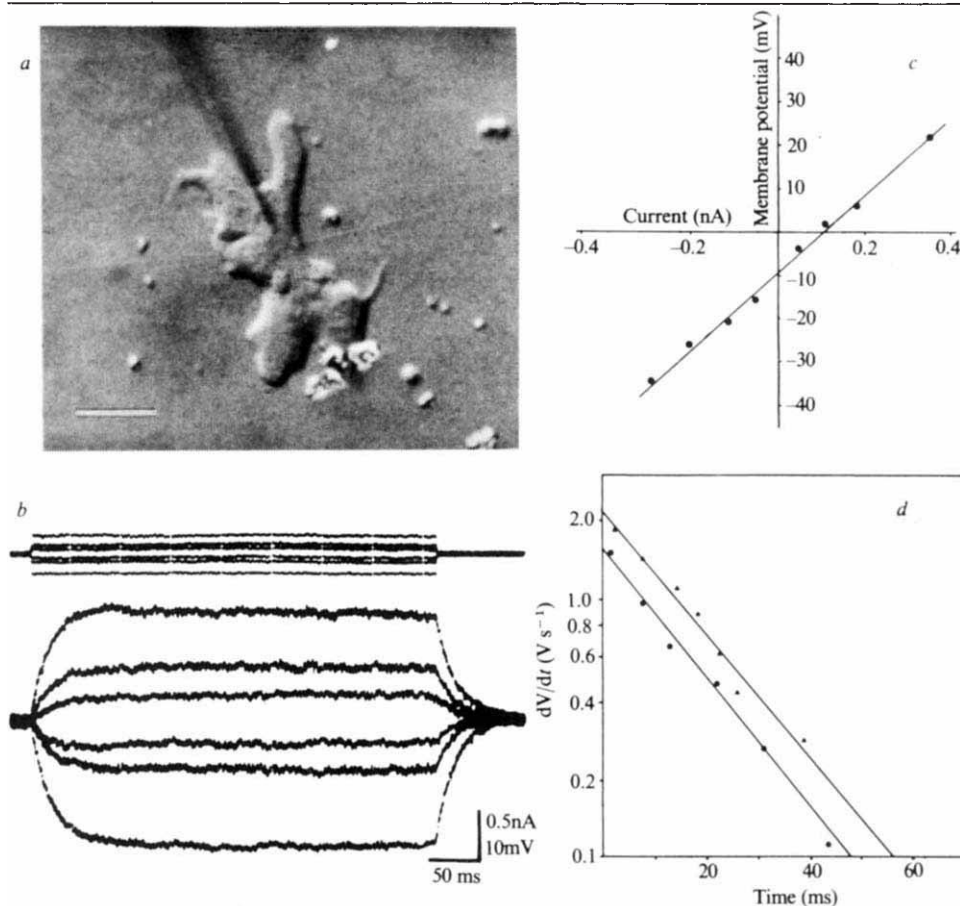


Fig. 1 Morphology and passive membrane properties of an isolated cone horizontal cell from the catfish retina. These cells were prepared by incubating isolated retinas with Ca^{2+} -free Ringer's solution supplemented with 5 mM EGTA and 1 mg ml⁻¹ papain (Worthington) at room temperature for 60–90 min. The retinas were then washed with Ca^{2+} -free Ringer's solution containing 1 mg ml⁻¹ bovine serum albumin and transferred to a 15-ml conical tube. Isolated cells were obtained by gently pipetting the retinas through a wide-bore Pasteur pipette in 2 ml of a medium containing 90% 0.3 M sucrose and 10% Ringer's solution. The cell suspension was kept on ice for 2 min before the top 1 ml was transferred to 9 ml oxygenated Ringer's solution (120 mM NaCl, 4 mM KCl, 2 mM NaHCO_3 , 5 mM CaCl_2 , 5 mM glucose, 15 mM HEPES buffer, pH 7.7) supplemented with 10 $\mu\text{g ml}^{-1}$ deoxyribonuclease². *a*, Photomicrograph of a cone horizontal cell before impalement, with a microelectrode above. The surface area of this cell was calculated to be 16,000 μm^2 . *b*, Voltage transients (bottom traces) recorded intracellularly from the cell shown in *a* in response to 400-ms constant current pulses (top traces). The cell was at its resting potential of -10 mV. *c*, Steady-state V - I plot from data shown in *b*. The slope indicates an input resistance of 89.6 M Ω . *d*, The first derivative of two voltage transients was determined graphically as described elsewhere^{7,13} and its log plotted against time. ●, Data points from a 10-mV depolarizing voltage transient; ▲, data points from a 15-mV hyperpolarizing transient. The data points can be fitted by single exponentials, suggesting isopotentiality for this cell. The average time constant was 17.7 ms.

In our electrophysiological studies, we started by isolating cone horizontal cells from retinas of the channel catfish for three reasons. First, the catfish retina contains only one type of cone (maximum absorbance 620 nm) and therefore only one type of cone horizontal cell that can be easily distinguished, even in isolation, from the more stellate rod horizontal cells⁸. This ensures that our recordings are from a homogeneous population of horizontal cells. Second, the somata of these horizontal cells, which are probably GABAergic⁹, receive input from dopaminergic terminals (D.M.K.L., in preparation). Therefore this preparation has a potential use in analyses of dopaminergic responses in a neurone from the vertebrate central nervous system. Third, these cells are relatively large in size (50–100 μm in diameter), making it possible to use low-impedance microelectrodes for voltage-clamp studies.

Catfish retinas were dissociated into single cells using a procedure described elsewhere². For electrophysiological studies, 0.2 ml of the isolated cell suspension was placed on a depression slide and the cells were allowed to sediment for 5 min. Under Nomarski interference optics, cone horizontal cells were impaled with 30–50 M Ω KCl-filled micropipettes. The single-electrode clamp system (SEC) used for current-clamp of these cells is described in detail elsewhere^{6,7}. Briefly, the SEC switches a single microelectrode rapidly (3 kHz) between current-passing and voltage-recording modes. The technique allows for the passage of larger transmembrane currents, and for a more accurate measurement of the membrane potential during current passage, than is possible using standard Wheatstone bridge arrangements. Each cell was photographed before intracellular recording. Surface area measurements were determined by projecting an enlarged image of the cell onto graph paper. The cells were assumed to be flat disks with a thickness of 10–15 μm , which was estimated from the depth of focus.

The results shown here are based on intracellular recordings of >92 cone horizontal cells. Figure 1*a* is a photomicrograph of a cone horizontal cell from which an intracellular recording was

made. Cone horizontal cells are easily identified by two criteria: they are the largest cells in the retina^{2,8} and even in isolation, they still have the amoeboid shapes characteristic of horizontal cells in the intact retina and can be easily distinguished from other isolated cells^{2,8}. Figure 1*b–d* gives data on the passive membrane properties of the cell shown in Fig. 1*a*. The isolated cone horizontal cells used here usually had an amoeboid shape with numerous fine processes. We generally chose to impale the larger cells with the simplest morphology that could be clearly and unequivocally identified, but that were free of long processes. For the cell in Fig. 1*a*, 400-ms constant current pulses were applied and the resulting voltage transients recorded (see Fig. 1*b*). Figure 1*c* shows a plot of the steady-state voltage-current (V - I) relationship, obtained from transients such as those shown in Fig. 1*b*. In this particular cell, the V - I plot is linear over the voltage range -40 to +20 mV. However, few cells showed completely linear V - I plots over a wide voltage range. We typically measured a decrease in input resistance with large depolarizations or hyperpolarizations.

To calculate the specific membrane capacity of these cells, it was necessary to determine accurately the membrane time constant. Figure 1*d* shows the logarithm of the first derivative of two low-amplitude voltage transients (from Fig. 1*b*), one depolarizing and the other hyperpolarizing, plotted against time. The data points fit well to single exponentials with time constants of 17.3 and 18.2 ms, respectively. As we could detect only a single exponential in the low-amplitude voltage transients, the cell behaved as a single isopotential compartment, composed of a parallel resistor and capacitor⁷. Moreover, the time constants calculated from the small depolarizing (10 mV) and hyperpolarizing (15 mV) transients are similar, which supports our assumption that the membrane resistivity is effectively ohmic in this narrow voltage range.

With the information from Fig. 1, the specific membrane properties of this cell can be calculated. The surface area (A) was estimated from Fig. 1*a* as 16,000 μm^2 . The input resistance (R_N) was 89.6 M Ω , and the average membrane time constant (τ_m) was

17.7 ms. These values yield a specific membrane resistivity (R_m) of $14,330 \Omega \text{ cm}^2$ and a specific membrane capacity (C_m) of $1.2 \mu\text{F cm}^{-2}$.

In most cells, depolarizing current pulses revealed a regenerative voltage response as defined by an increasing positive slope in the membrane potential transient. An example of such a regenerative response is shown in Fig. 2. The resting membrane potentials of impaled cells from different preparations were quite variable, ranging from -10 to -100 mV. However, as most isolated cells in a given preparation had similar resting potentials, all with large input resistances, we believe that these variations in resting potentials are not the result of damage due to the microelectrode, but are most likely due to variations in the retinæ used or in the isolation procedures. Cells with low resting potentials generally showed regenerative voltage responses if a steady hyperpolarizing current was applied before the depolarizing pulses. The hyperpolarization probably reduced delayed rectification and closed the voltage-dependent channels that are responsible for the regenerative response. The cell shown in Fig. 2 had a resting membrane potential of -25 mV, but the voltage transients were obtained when the cell was hyperpolarized to -65 mV. The passive properties of this cell were $R_N = 250 \text{ M}\Omega$ at -25 mV, $A = 3,740 \mu\text{m}^2$, $R_m = 9,350 \Omega \text{ cm}^2$, $\tau_m = 9.6$ ms and $C_m = 1.0 \mu\text{F cm}^{-2}$.

The addition of tetrodotoxin (TTX, $10 \mu\text{g ml}^{-1}$) to the bathing solution of these cells had no apparent effect on the regenerative responses. To test whether Ca^{2+} might be involved in the regenerative responses, we applied CoCl_2 , MnCl_2 or NiCl_2 by pressure injection to the bathing solution; all three of these divalent cations have been reported to block inward Ca^{2+} currents in various systems^{10,11}. We found a rapid and complete blockade of all regenerative voltage responses with the application of any one of these divalent cations at a final concentration of ~ 10 mM, although Ni^{2+} seemed to be the most effective. Figure 3a shows the regenerative voltage response from a cell bathed in $10 \mu\text{g ml}^{-1}$ TTX. The transients in Fig. 3b were taken a few minutes after applying 10 mM Ni^{2+} ; the regenerative response could no longer be elicited. Similar results were obtained in six other cells. These experiments demonstrate that the regenerative voltage responses recorded in isolated cone horizontal cells are TTX-insensitive and probably Ca^{2+} -dependent.

The results from a number of isolated cells show that they have an average specific membrane capacity of $1.3 \pm 0.3 \mu\text{F cm}^{-2}$ (mean $\pm$ s.d., $n = 7$). A C_m of $1 \mu\text{F cm}^{-2}$ has been suggested to be a biological constant, as this value has been found for all carefully studied cells in both vertebrate and invertebrate nervous systems^{12,13}. Our results agree with this hypothesis. However, it should be emphasized that our data may have several sources of error. First, the sample of cells is limited and biased towards larger cells with the simplest morphologies. Second, the assumption of uniform thickness for each cell and the inaccuracy associated with measuring depth could increase our surface area estimate by 10–20%. Third, the measurement

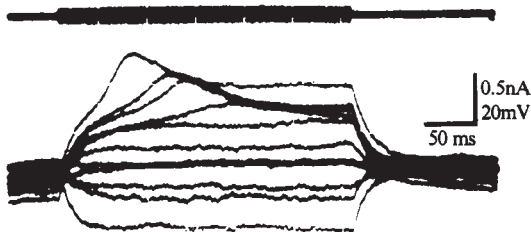


Fig. 2 Regenerative voltage responses from a cone horizontal cell. The resting potential of this cell was -25 mV. With the application of steady hyperpolarizing current, which brought the membrane potential to -65 mV, regenerative responses could be obtained in response to 300-ms constant current pulses (top traces). The passive membrane properties of this cell, determined as described in Fig. 1 legend, were: $R_N = 250 \text{ M}\Omega$, $A = 3,740 \mu\text{m}^2$, $R_m = 9,350 \Omega \text{ cm}^2$, $\tau_m = 9.6$ ms and $C_m = 1.0 \mu\text{F cm}^{-2}$.

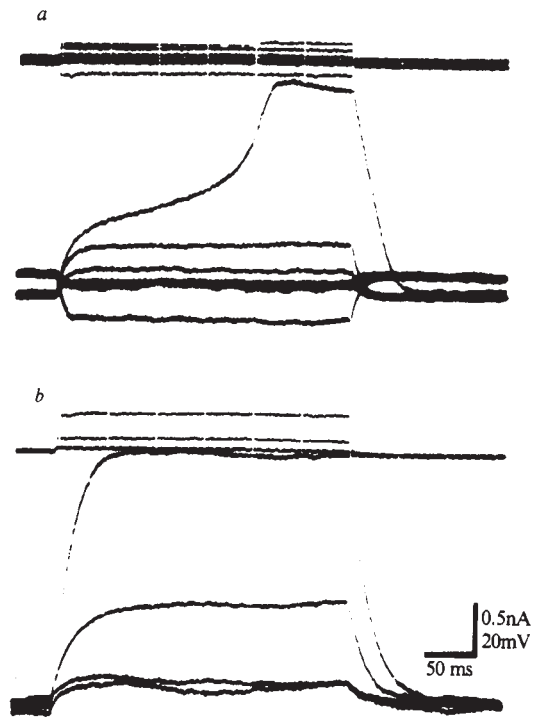


Fig. 3 Regenerative voltage responses blocked by Ni^{2+} . *a*, A regenerative voltage transient in response to a 300-ms constant current pulse. The resting potential of this cell was ~ -50 mV. Steady hyperpolarizing current brought the cell to -90 mV, from which the pulses shown were given. The cell was bathed in $10 \mu\text{g ml}^{-1}$ TTX. *b*, The same cell several minutes after the application of 10 mM Ni^{2+} . No regenerative responses could be recorded with the cell at -90 mV. Steady depolarizing and hyperpolarizing currents were passed to bring the cell to several different membrane potentials. The application of depolarizing current pulses from any of these different membrane potentials could not elicit a regenerative voltage response.

of the membrane time constant could be influenced by voltage-dependent currents. This last source of error would be most critical in those cells with the largest specific membrane resistivities.

With the exception of an apparent lack of TTX-sensitive sodium channels, the active and passive membrane properties of cone horizontal cells seem to be similar to those described in various neurones from the central and peripheral nervous systems of both vertebrate and invertebrate species. Our results suggest that catfish cone horizontal cells, in isolation, can show regenerative behaviour that is probably Ca^{2+} -dependent. The leak current *in vivo*, caused by the extensive electrotonic couplings among horizontal cells, might be expected to subtract from a voltage-dependent inward current and prevent the appearance of regenerative voltage responses. Accurate estimates of membrane resistivity are also difficult to make in the intact retina because of these electrical couplings between horizontal cells¹⁴. For example, the input resistances of horizontal cells in intact retinæ were reported to be of the order of $500 \text{ k}\Omega$ (ref. 14), a value about 200 times lower than that found in isolated horizontal cells. For these reasons, the ability to record from isolated cells is crucial for the study of the biophysical properties of these neurones.

We thank Drs A. M. Brown and T. H. Brown for critical reading of the manuscript and Ms P. Glazebrook and Mr L. M. F. Lam for technical assistance. This work was supported by NIH grants NS15772, NS11535 (to D. J.) and EY02423 (to D.M.K.L.) and by a grant from the Retina Research Foundation (Houston). D.M.K.L. is the recipient of a Research Career Development Award from the US National Eye Institute and an Olga K. Weiss Scholar Award from Research to Prevent Blindness, Inc., New York.

Received 19 January; accepted 18 May 1981.

1. Lam, D. M.K. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1987 (1972).
2. Lam, D. M.K. *Nature* **254**, 345–347 (1975); *Cold Spring Harb. Symp. quant. Biol.* **40**, 571–579 (1976).

3. Drujan, B. D. & Svaetichin, G. *Vision Res.* **12**, 1777-1784 (1972).
4. Kaneko, A., Lam, D. M. K. & Wiesel, T. N. *Brain Res.* **105**, 567-572 (1976).
5. Bader, C. R., Macleish, P. R. & Schwartz, E. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3507-3511 (1978).
6. Johnston, D., Hablitz, J. J. & Wilson, W. A. *Nature* **286**, 391-393 (1980).
7. Johnston, D. *Cell. molec. Neurobiol.* **1**, 41-55 (1981).
8. Naka, K.-I. & Caraway, N. R. G. *J. Neurophysiol.* **38**, 53-71 (1975).
9. Lam, D. M. K., Lasater, E. M. & Naka, K.-I. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6310-6313 (1978).
10. Hagiwara, S. in *Membrane—A Series of Advances* Vol. 3 (ed. Eisenman, G.) 359-381 (Dekker, New York, 1975).
11. Akaike, N., Lee, K. S. & Brown, A. M. *J. gen. Physiol.* **71**, 509-531 (1978).
12. Cole, K. S. *Membranes, Ions and Impulses* (University of California Press, Berkeley, 1968).
13. Brown, T. H., Perkel, D. H., Norris, J. C. & Peacock, J. H. *J. Neurophysiol.* **45**, 1-15 (1981).
14. Kaneko, A. *J. Physiol., Lond.* **213**, 95-105 (1971).

Inhibition of basophil histamine release by anti-inflammatory steroids

Robert P. Schleimer, Lawrence M. Lichtenstein & Elizabeth Gillespie

Department of Medicine, Division of Clinical Immunology, Johns Hopkins University School of Medicine at The Good Samaritan Hospital, Baltimore, Maryland 21239, USA

The release of histamine from human leukocytes is a useful *in vitro* model for studying allergic disease¹. Many of the drugs used in the treatment of allergy and asthma (for example, isoprenaline and theophylline) are effective inhibitors of *in vitro* histamine release². However, the anti-inflammatory steroids have not been found to inhibit the *in vitro* release of histamine from mast cells or basophils activated by immunological stimuli³⁻⁶. In view of the fact that the *in vivo* anti-allergic effects of steroids occur only 12-24 h after administration⁷, we have re-examined the effects of these drugs on IgE-mediated histamine release from human basophils after prolonged incubations. We report here a time-dependent inhibition of histamine release from human basophils by glucocorticosteroids. The relative activity of a series of steroids as inhibitors of histamine release correlates well with the relative *in vivo* anti-inflammatory activity of the same compounds. These results suggest that the basophil may be an *in vivo* target of the anti-inflammatory steroids and that inhibition of this cell type may account for some of the activity of steroids in preventing inflammation.

Human leukocytes were prepared from the blood of normal volunteers by dextran sedimentation as described elsewhere⁸ except that sterile conditions were used. Cells were cultured in the presence or absence of steroids for up to 24 h, as described in Fig. 1 legend, after which they were washed and challenged with either anti-IgE or buffer control. Dexamethasone acetate, at concentrations up to 10^{-6} M, did not reduce the total cellular histamine content in incubations of 24 h. Total leukocyte viability, as assessed by Trypan blue exclusion, was always greater than 70% and usually greater than 80% after 24-h culture with or without up to 10^{-6} M dexamethasone acetate. In initial experiments, the drug in use was present during the wash and challenge procedures. However, it was determined that its omission during wash and challenge steps did not affect the extent of inhibition of histamine release, and it was therefore not included in most of the experiments described here.

The time course of onset of dexamethasone-induced inhibition of histamine release is shown in Fig. 1. Control release in the absence of steroid was 28%. Preincubation of cells with dexamethasone acetate for up to 2 h produced no inhibition of histamine release during the subsequent 45-min challenge. However, inhibition was noted after 4 h of incubation, and increased linearly for 24 h. Similar results were obtained with anti-IgE in three experiments. In 21 experiments using a 24-h incubation, histamine release in the control groups, stimulated with $0.1 \mu\text{g ml}^{-1}$ anti-IgE, was $34 \pm 3.3\%$ (mean $\pm$ s.e.m.) whereas that from cells exposed to 10^{-7} M dexamethasone acetate was $12.9 \pm 1.4\%$. In two experiments, using the cells of

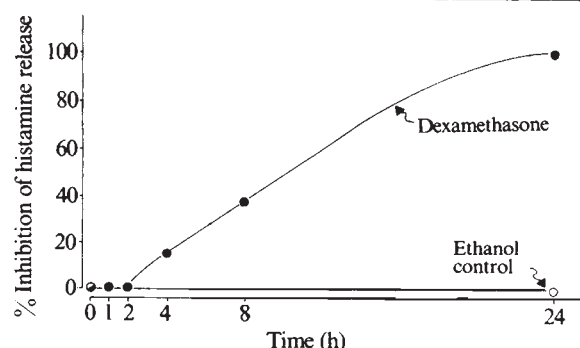


Fig. 1 Human leukocytes were collected from adult volunteers, sedimented in sterile dextran/EDTA, washed twice in sterile saline and cultured in 2-ml aliquots in multi-well tissue culture plates for 24 h in RPMI 1640 medium supplemented with penicillin/streptomycin, glutamine, 25 mM HEPES and 7% autologous serum. Dexamethasone acetate in ethanol at 10^{-2} M, or ethanol alone, was diluted in medium to a final concentration of 10^{-7} M and added at various times so that total duration of exposure to the drug was as indicated. At the end of the culture period, cells were washed in PIPES-buffered saline, pH 7.3, containing 110 mM NaCl, 5 mM KCl and 25 mM PIPES, and resuspended in PIPES-buffered saline containing 1 mM CaCl_2 and 1 mM MgCl_2 . Cells were then challenged with either $0.1 \mu\text{g ml}^{-1}$ anti-IgE (a gift of Dr K. Ishizaka) or buffer control (spontaneous release) at 37°C for 45 min. Histamine released into the supernatant was measured using an automated fluorometric technique²¹. Per cent histamine release was calculated as the fraction ($\times 100$) of total cellular histamine released into the challenge medium after subtraction of the spontaneous release ($<5\%$). Per cent inhibition of histamine release was calculated as the reduction (expressed on a percentage basis) compared with control release induced by drug treatment.

ragweed-sensitive donors, histamine release stimulated by anti-IgE was inhibited by over 80% after a 24-h incubation with 10^{-7} M dexamethasone acetate.

To ascertain whether the *in vitro* activity of steroids parallels their *in vivo* effects, several steroids were tested for activity as inhibitors of basophil histamine release using 24-h incubation periods. The inhibition of histamine release by four anti-inflammatory steroids showed a rank order of triamcinolone acetate > dexamethasone acetate > 9α -fluorocortisone > hydrocortisone (Fig. 2). This order of potency is consistent with the anti-inflammatory potency of these compounds in man⁹, and the IC_{50} of these and other steroids shown in Table 1 are similar to their binding constants for the steroid receptor^{10,11}. The sex steroids, β -oestradiol, testosterone and progesterone, as well as the cortisone metabolite, tetrahydrocortisone, were essentially inactive as inhibitors of histamine release. Prednisone, a cortisone analogue that is active *in vivo*, was inactive as an inhibitor of histamine release, whereas prednisolone, in two experiments, was an effective inhibitor with an approximate IC_{50} of 2.2×10^{-8} M. This observation is consistent with the hypothesis that prednisolone, the major metabolite of pred-

Table 1 Potency of various steroids as inhibitors of histamine release

Compound	IC_{50} (M)	n	Relative potency
Triamcinolone acetate	3.1×10^{-9}	6	48
Dexamethasone acetate	7.8×10^{-9}	6	19
Prednisolone	2.2×10^{-8}	2	6.8
9α -Fluorocortisone	4.5×10^{-8}	4	3.3
Hydrocortisone	1.5×10^{-7}	8	1
Progesterone	$>10^{-5}$	1	<0.01
β -Oestradiol	$>10^{-5}$	2	<0.01
Testosterone	$>10^{-5}$	2	<0.01
Prednisone	$>10^{-5}$	2	<0.01
Tetrahydrocortisone	$>10^{-5}$	3	<0.01

IC_{50} , molar concentration of drug that inhibits histamine release by 50%. n, Number of experiments from which data are derived. Potency is relative to hydrocortisone.

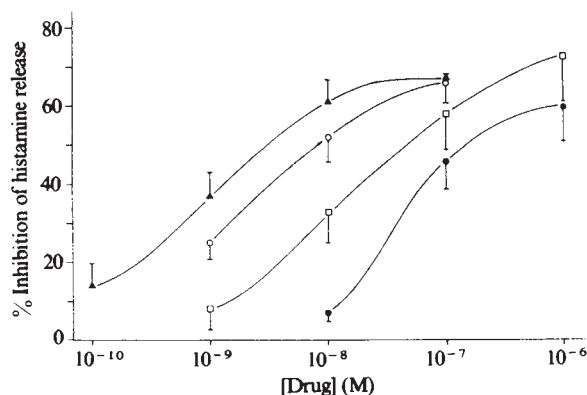


Fig. 2 Inhibition of basophil histamine release by triamcinolone acetate (▲), dexamethasone acetate (○), 9α-fluorocortisone (□) and hydrocortisone (●). All steroids were dissolved in ethanol at 10⁻² M and subsequently diluted in medium. Assay and calculations were performed as described in Fig. 1 legend and in the text. Data shown represent the mean ± s.e.m. inhibition of at least four experiments for each drug tested (see Table 1).

nisone, is responsible for the *in vivo* activity of prednisone^{12,13}. The relative and absolute potencies of the steroids tested are similar to their potencies as inhibitors of arachidonic acid release from guinea pig lung¹⁴ and transformed mouse fibroblasts¹⁵, as well as their potencies as inhibitors of lymphocyte mitogenesis¹⁶ and macrophage plasminogen activator release¹⁷. Note that results similar to ours have recently been obtained using mouse peritoneal mast cells (M. Daëron, A. Sterk and T. Ishizaka, personal communication).

The mechanism of the steroid effect is unclear, but the observation that basophil activity is inhibited by glucocorticosteroids raises several important questions. First, it becomes necessary to determine to what extent a steroid-mediated inhibition of basophil activity plays a part in the abrogation of those delayed-type hypersensitivities which involve basophils^{18,19}. Second, as hydrocortisone inhibits histamine release at concentrations which are found naturally in the circulation⁹, it becomes necessary to determine what role circulating glucocorticoids have in the regulation of basophil activity. Third, questions concerning the biochemical mechanism by which the steroids inhibit histamine release can be approached using a recent technique for obtaining human basophils at high purity²⁰. Finally, it becomes necessary to ask what role the inhibition of basophil activity has in the glucocorticosteroid alleviation of allergic disease and asthma.

This work was supported by NIH grants AI 07290, AI 14959, AI 17301 and HL 23586. Publication no. 436 of the O'Neill Laboratories, The Good Samaritan Hospital.

Received 12 March; accepted 26 May 1981.

- Lichtenstein, L. M., Marone, G., Thomas, L. L. & Malveaux, F. J. *invest. Derm.* **71**, 65–69 (1978).
- Bourne, H. R. *et al. Science* **184**, 19–28 (1974).
- Norn, S. *Acta pharmac. tox.* **22**, 369–378 (1965).
- Schmutzler, W. & Freundt, G. P. *Int. Archs Allergy appl. Immun.* **49**, 209–212 (1975).
- Thomas, R. O. & Whittle, B. J. R. *Br. J. Pharmac.* **57**, 474P–475P (1976).
- Lewis, G. P. & Whittle, B. J. R. *Br. J. Pharmac.* **61**, 229–235 (1977).
- Liddle, G. W. *Clin. Pharmac. Ther.* **2**, 615–635 (1961).
- Lichtenstein, L. M. & Osler, A. G. *J. exp. Med.* **120**, 507–530 (1964).
- Gilman, A. G., Goodman, L. S. & Gilman, A. *The Pharmacological Basis of Therapeutics* 6th edn (Macmillan, New York, 1980).
- Ballard, P. L., Carter, J. P., Graham, B. S. & Baxter, J. D. *J. clin. Endocr. Metab.* **41**, 290–304 (1975).
- Dausse, J. P. *et al. Molec. Pharmac.* **13**, 948–955 (1977).
- Meikle, A. W., Weed, J. A. & Tyler, F. H. *J. clin. Endocr. Metab.* **41**, 717–721 (1975).
- Peterson, R. E., Pierce, C. E., Wyngaarden, J. B., Bunim, J. J. & Brodie, B. B. *J. clin. Invest.* **37**, 1301–1312 (1957).
- Blackwell, G. J., Flower, R. J., Nijkamp, F. P. & Vane, J. R. *Br. J. Pharmac.* **62**, 79–89 (1978).
- Tam, S., Hong, S. L. & Levine, L. *J. Pharmac. exp. Ther.* **203**, 162–168 (1977).
- Cantrill, H. L., Waltman, S. R., Palmberg, P. F., Zink, H. A. & Becker, B. J. *clin. Endocr. Metab.* **40**, 1073–1077 (1975).
- Vassalli, J. D., Hamilton, J. & Reich, E. *Cell* **8**, 271–281 (1976).
- Dvorak, A. M., Mihm, M. C. & Dvorak, H. F. *J. Immun.* **116**, 687–695 (1976).
- Askenase, P. W., Haynes, D. & Hayden, B. J. *J. Immun.* **117**, 216–224 (1976).
- MacGlashan, D. W. & Lichtenstein, L. M. *J. Immun.* **124**, 2519–2521 (1980).
- Siraganian, R. P. *Analyt. Biochem.* **57**, 383–388 (1974).

Lipoxygenase products modulate histamine release in human basophils

Stephen P. Peters, Marvin I. Siegel,
Anne Kagey-Sobotka & Lawrence M. Lichtenstein

The Johns Hopkins University School of Medicine, Department of Medicine, Divisions of Clinical Immunology and Pulmonary Medicine, Baltimore, Maryland 21239 and Burroughs Wellcome Co., Research Triangle Park, North Carolina 27709, USA

Arachidonic acid released from cellular phospholipids is metabolized by two major enzyme systems¹. The cyclooxygenase pathways are responsible for the production of prostaglandins, thromboxanes and prostacyclin while the lipoxygenase pathways form a series of monoperoxy- and monohydroxyeicosatetraenoic acids. The 5-lipoxygenase pathway eventually leads to the production of both leukotriene B₄, one of the most potent chemotactic molecules known, and SRS-A (slow-reacting substance of anaphylaxis—leukotrienes C and D). Our earlier studies of the lipoxygenase enzyme(s) suggested that a product(s) of lipoxygenase(s) is both necessary for histamine release from human basophils and blunts endogenous control mechanisms^{2–4}. To test this hypothesis, we have studied the effect of exogenously added products of lipoxygenase pathways on antigen-induced histamine release from human basophilic leukocytes *in vitro*. The present report demonstrates that 5-hydroperoxyeicosatetraenoic acid (5-HPETE), at sub-micromolar concentrations, causes both a dose-dependent enhancement of histamine release and a reversal of the inhibition of histamine release caused by hormones and other agonists which act via adenylate cyclase. The reduced product of 5-HPETE, 5-hydroxyeicosatetraenoic acid (5-HETE) has both activities, but is approximately 3 to 10 times less potent.

As studies of the oxidation of arachidonic acid by cyclooxygenase and lipoxygenase pathways have been pursued, evidence has accumulated that products of the lipoxygenase pathways and other fatty acid hydroperoxides: (1) enhance anaphylactic mediator release from guinea pig lung^{5,6}, (2) are chemotactic for polymorphonuclear leukocytes^{7–14}, (3) can inhibit both leukotriene biosynthesis¹⁵ and platelet cyclooxygenase activity¹⁶, (4) enhance IgE-mediated release of histamine in rat mast cells¹⁷, (5) are incorporated into neutrophil phospholipids and triglycerides¹⁸ and induce degranulation of these cells¹⁹, and (6) may be involved in inflammatory and pyretic reactions *in vivo*²⁰.

The nonspecific lipoxygenase inhibitor, 5,8,11,14-eicosatetraenoic acid, and 5,8,11-eicosatrienoic acid, a relatively specific lipoxygenase inhibitor, inhibit basophil histamine release caused by antigen–IgE interactions, C5a peptide, fMet peptides and the ionophore, A23187, in a complete and dose-dependent fashion^{2,3}. In addition, arachidonic acid and pharmacologic concentrations of indomethacin and other non-steroidal anti-inflammatory drugs both potentiate histamine release and reverse the inhibition of histamine release caused by drugs and hormones that act via adenylate cyclase, while failing to affect the action of drugs which increase cyclic AMP levels by other mechanisms^{2–4}. These data, together with the finding that the phospholipase A₂ inhibitor, *p*-bromophenacyl bromide, also causes a complete and dose-dependent inhibition of histamine release⁴, led to the hypothesis that a lipoxygenase product(s) of arachidonic acid, generated by phospholipase A₂, was both obligatory for histamine release and negatively modulated the signal generated by receptors linked to adenylate cyclase.

Figure 1 shows the dose-response relationships for the enhancement of histamine release by 5-HPETE and 5-HETE. 5-HPETE had half-maximal effect at 0.10 μM, and a peak effect

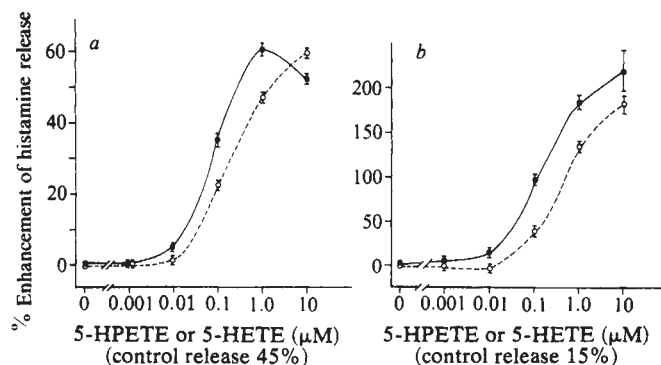


Fig. 1 Enhancement of basophil histamine release by 5-HPETE and 5-HETE. Peripheral blood leukocytes from atopic adult volunteers who had given informed consent were isolated by dextran sedimentation of venous blood anticoagulated with EDTA as previously described². The histamine release assay was performed in a total volume of 0.5 ml in buffer containing 25 mM PIPES, 110 mM NaCl, 5 mM KCl, 0.003% human serum albumin, 1.0 mM MgCl₂ and 1.0 mM CaCl₂. Incubations were carried out for 45 min at 37 °C after a 5-min preincubation of cells with drug or lipoxygenase product before addition of antigen, either ragweed antigen E or rye grass, group I. The lipoxygenase product, 5-HPETE, was prepared as described by Porter *et al.*²²; 5-HETE was prepared from 5-HPETE by reduction with NaBH₄. These lipoxygenase products were added to 12 × 75-mm polypropylene reaction tubes as an ether solution which was evaporated under a stream of nitrogen immediately before use. Leukotrienes C and D were a gift from Dr Joseph Rokach of Merck, Sharp and Dohme and stored frozen as an aqueous solution at -70 °C. Dimaprit was a gift of Dr M. Parsons of Smith Kline and French. Histamine released was quantified by fluorometric assay as previously described^{23,24} and corrected for spontaneous release, typically <5% of the total histamine content. Experiments were carried out in duplicate or triplicate with cells from 3–10 atopic donors. None of the lipoxygenase products caused spontaneous histamine release. In the experiment above, assays were performed in triplicate with either 5-HPETE (●) or 5-HETE (○). Control histamine release was 45% (a) and 15% (b). Error bars indicate s.e.m.

at 1.0 μM. 5-HETE had a similar effect but was 60–80% as active as 5-HPETE in enhancing histamine release at a concentration of 1 μM. Examination of the linear portion of the enhancement dose–response curve shows that it requires 3 to 10 times more 5-HETE than 5-HPETE to cause an equivalent enhancement of histamine release. The enhancement observed is, of course, a function of the magnitude of the control histamine release. In a series of 20 experiments, the enhancement ranged from 15 to 60% when histamine release was 40–70% of total cellular histamine (Fig. 1a), and from 190 to 340% when histamine release was 10–15% of total cellular histamine (Fig. 1b). In a series of six experiments, 11-HPETE and 15-HPETE were 100 ± 40 times less active than 5-HPETE in enhancing release. The stable hydroperoxide, 13-hydroperoxylinoleic acid, showed no enhancement at this concentration in four of six experiments. In six experiments, other products of the 5-lipoxygenase pathway, leukotrienes C and D, at concentrations of 0.5–3 μM, caused no consistent enhancement (or inhibition) of antigen-induced histamine release. These results are in contrast to the observations of Schellenberg and co-workers, who showed differential effects for leukotriene C and D on 48/80-induced histamine release in rat mast cells²¹. Other lipoxygenase products, including the potent chemotactic molecule, leukotriene B^{12–14}, were not available for evaluation.

Figure 2 demonstrates the previously reported specific reversal by indomethacin of the inhibition caused by the histamine type 2 agonist, dimaprit². The activity of 5-HPETE is essentially the same as that of indomethacin. 5-HETE also reversed the inhibition caused by dimaprit but to a lesser extent (data not shown). Similar reversal was seen in five of five experiments with 5-HPETE and four of five experiments with 5-HETE. In simul-

taneous experiments, neither indomethacin (as previously reported) nor 5-HPETE reversed the inhibition caused by dibutyl cyclic AMP or isobutyl methylxanthine, agonists which increase cyclic AMP levels by mechanisms other than activation of adenylate cyclase (that is, directly or by phosphodiesterase inhibition). Thus, the effect of 5-HPETE in specifically reversing the inhibition of histamine release caused by agonists which act via adenylate cyclase is the same as that previously reported for indomethacin and arachidonic acid.

The mechanism by which antigen–IgE interactions cause mediator release from basophils and mast cells has been extensively studied^{2–4,17–19}. The bulk of the evidence suggests that stimulation of the basophil occurs by cross-linking cell-surface IgE molecules followed by phospholipase A₂ activation leading to the production of arachidonic acid which is then metabolized by the lipoxygenase pathway(s) leading to histamine release. The present study provides direct evidence that specific 5-lipoxygenase products (5-HPETE and 5-HETE) are able to modulate histamine release in human basophils; they can both enhance antigen-induced histamine release and specifically reverse the inhibition caused by agents that increase cyclic AMP levels via an effect on adenylate cyclase. Products of the 11- and 15-lipoxygenase pathways are much less active in this regard and leukotrienes C and D are inactive. The naturally occurring hydroperoxy derivatives of arachidonic acid are relatively unstable in aqueous solutions, being converted, usually within minutes, to the corresponding hydroxy compound. The present report demonstrates that both the relatively unstable 5-hydroperoxy derivative of arachidonic acid and the stable 5-hydroxy-eicosatetraenoic acid have the ability to modulate histamine release in human basophils. Whether further metabolism of these compounds occurs during the *in vitro* incubation is unknown. Our interpretation assumes that exogenously added material will have a similar effect to endogenously generated products.

The mechanism by which lipoxygenase products enhance histamine release is unknown, although it is tempting to speculate that these products may be incorporated into membrane phospholipids and then induce degranulation, analogous to the mechanism proposed for human neutrophils^{18,19}. The ability of other lipoxygenase products to modulate enzymes involved in inflammatory reactions, including the inhibition of platelet cyclooxygenase activity¹⁶ and the inhibition of the biosynthesis of leukotrienes¹⁵, argues strongly for the importance of these compounds in inflammatory processes. Whether other lipoxygenase products, particularly leukotriene B, are capable of modulating histamine release in human basophils remains to be tested.

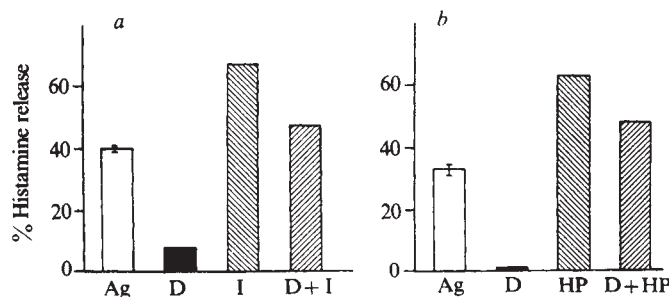


Fig. 2 Reversal of dimaprit inhibition of histamine release by indomethacin (a) and 5-HPETE (b). Cells were preincubated for 5 min with dimaprit (2 μM), indomethacin (2 μM), 5-HPETE (1 μM) or combinations of the above for 5 min before addition of antigen. Assays containing drug or lipoxygenase products were performed in duplicate; in general, replicate assays varied by <10%. Ag, antigen controls; D, dimaprit; I, indomethacin; HP, 5-HPETE. Error bars indicate range of histamine release for antigen controls.

We thank Ms Renee Daiuta and Mrs LaVerne Grove for technical assistance. This work was supported by NIH grants AI 07290 and HL 23586. S.P.P. is a postdoctoral fellow supported by training grant HL 07199 from Johns Hopkins University School of Medicine. This is publication no. 437 from the O'Neill Laboratories, The Good Samaritan Hospital.

Received 24 March; accepted 27 May 1981.

1. Samuelsson, B., Gränström, E., Green, K., Hamberg, M. & Hammarström, S. A. Rev. *Biochem.* **44**, 669-695 (1975).
2. Marone, G., Kagey-Sobotka A. & Lichtenstein L. M. *J. Immun.* **123**, 1664-1667 (1979).
3. Marone, G., Hammarström S. & Lichtenstein, L. M. *Clin. Immun. Immunopath.* **17**, 117-122 (1980).
4. Marone, G., MacGlashan, D. W. Jr, Kagey-Sobotka, A. & Lichtenstein, L. M. in *Advances in Allergy and Immunology* (eds Oehling, A. et al.) 147-154 (Pergamon, New York, 1980).
5. Adcock, J. J., Garland, L. G., Moncada, S. & Salmon, J. A. *Prostaglandins* **16**, 163-177 (1978).
6. Adcock, J. J., Garland, L. G., Moncada, S. & Salmon, J. A. *Prostaglandins* **16**, 179-187 (1978).
7. Turner, S. R., Tainer, J. A. & Lynn, W. S. *Nature* **257**, 680-681 (1975).
8. Goetzl, E. J., Woods, J. M. & Gorman, R. R. *J. clin. Invest.* **59**, 179-183 (1977).
9. Goetzl, E. J., Weller, P. F. & Sun, F. F. *J. Immun.* **124**, 926-933 (1980).
10. Goetzl, E. J. & Sun, F. F. *J. exp. Med.* **150**, 406-411 (1979).
11. Goetzl, E. J., Hill, H. R. & Gorman, R. R. *Prostaglandins* **19**, 71-85 (1980).
12. Ford-Hutchinson, A. W., Bray, M. A., Doig, M. V., Shipley, M. E. & Smith, M. J. H. *Nature* **286**, 264-265 (1980).
13. Palmer, R. M. J., Stepney, R. J., Higgs, G. A. & Eakins, K. E. *Prostaglandins* **20**, 411-418 (1980).
14. Goetzl, E. J. & Pickett, W. C. *J. Immun.* **125**, 1789-1791 (1980).
15. Vanderhoek, J. Y., Bryant, R. W. & Bailey, J. M. *J. biol. Chem.* **255**, 10064-10065 (1980).
16. Siegel, M. A., McConnell, R. J., Abrahams, S. L., Porter, N. A. & Cuatrecasas, P. *Biochem. biophys. Res. Commun.* **89**, 1273-1280 (1979).
17. Stenson, W. F., Parker, C. W. & Sullivan, T. J. *Biochem. biophys. Res. Commun.* **96**, 1045-1052 (1980).
18. Stenson, W. F. & Parker, C. W. *Prostaglandins* **18**, 285-292 (1979).
19. Stenson, W. F. & Parker, C. W. *J. Immun.* **124**, 2100-2104 (1980).
20. Siegel, M. A., McConnell, R. T. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3774-3778 (1979).
21. Schellenberg, R. R., Johnston, M. E., Bach, M. K. & Hanna, C. J. *Fedn Proc.* **40**, 1014 (1981).
22. Porter, N. A., Legan, J. & Kontoyiannidou, V. *J. org. Chem.* **44**, 3177-3181 (1979).
23. Lichtenstein, L. M. & Osler, A. G. *J. exp. Med.* **120**, 507-530 (1964).
24. Siraganian, R. P. *Analyt. Biochem.* **57**, 383-394 (1974).

Type-specific protective immunity evoked by synthetic peptide of *Streptococcus pyogenes* M protein

**Edwin H. Beachey*†, Jerome M. Seyer‡,
James B. Dale*, W. Andrew Simpson*
& Andrew H. Kang*‡**

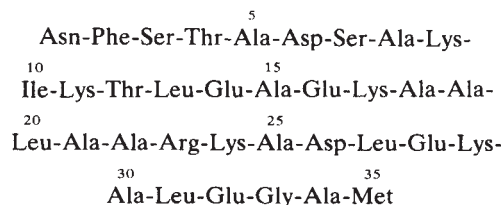
The Veterans Administration Medical Center and the Departments of *Medicine, †Microbiology and ‡Biochemistry, University of Tennessee School of Medicine, Memphis, Tennessee 38104, USA

The M protein on the surface of *Streptococcus pyogenes* cells enables the organisms to resist ingestion and killing by phagocytic cells in the blood of the non-immune host. In the immune host, type-specific antibodies against the M protein neutralize the antiphagocytic effect and allow the rapid elimination of any invading streptococci having the same serotype of M protein. Recent studies¹ in our laboratories of the primary structure of one of these M proteins (type 24) have indicated that the molecule is composed of repeating covalent structures, each of which contains protective antigenic determinants. These results suggest that only a small portion of the M protein molecule may be required to produce primary protective immunity against *S. pyogenes*. We now report that this notion is well founded. Our findings may have a bearing on the development of M-protein vaccines to protect against streptococcal infections causing acute rheumatic fever and rheumatic heart disease².

The mechanism whereby streptococcal infections give rise to complications such as rheumatic fever remains a mystery³. Because the sera of some patients with rheumatic fever show serological cross-reactivity between heart tissue antigens and certain streptococcal antigens³, it has been feared that immunization with intact M-protein vaccines may lead to

rheumatic heart disease. However, we have recently found that rabbits and mice immunized with cyanogen bromide fragments (CB6 or CB7) of type 24 M protein containing only 35 amino acid residues each developed opsonic and protective antibodies against type 24 streptococci⁴. One of these peptides (CB7) has now been chemically synthesized (S-CB7) and we report here protective immunity in laboratory animals induced by S-CB7 covalently linked to polylysine. In addition, we present evidence that one of the protective determinants resides in a peptide fragment of CB7 containing only 12 amino acid residues.

The CB7 polypeptide fragment of type 24 M protein and a dodecapeptide starting at residue 18 and ending at residue 29 (S18–29CB7) were synthesized⁵ and purified by reverse-phase HPLC on Ultrasphere ODS2 (Whatman) columns (Peninsula laboratories). The S18–29CB7 peptide overlaps two subpeptides derived from trypsin digestion of lysyl-blocked CB7 (see below). Amino acid analyses confirmed the identity of the synthetic peptides. Automated Edman degradation^{1,4} to the penultimate residue gave the amino acid sequence of S-CB7 as



This sequence is identical to that of the native CB7 fragment except that the COOH-terminal residue of S-CB7 is methionine, not homoserine⁴. The amino acid sequence of the S18-29CB7 peptide was found to be identical to the corresponding segment of CB7 except for the addition of a glycine tripeptide at the COOH terminal, which is used as a leash in the synthesis of the peptide (Penninsula Laboratories).

Three rabbits immunized with 25 nmol of S-CB7 emulsified in complete Freund's adjuvant (CFA)⁴ developed antibody titres at 6 weeks of 1:400, 1:1,280 and 1:25,600, respectively, as determined by enzyme-linked immunosorbent assays (ELISAs)^{2,6}. However, only the serum showing the highest ELISA titre was capable of opsonizing type 24 streptococci (data not shown). Therefore, three additional rabbits were immunized with 25 nmol of CB7 covalently conjugated to polylysine (molecular weight (MW) $\approx$ 35,000) and emulsified in CFA. The sera of all three rabbits showed good antibody responses as measured by ELISAs or opsonic antibody assays (Fig. 1). In bactericidal assays (see Fig. 1 legend) using types 5, 6 and 24 streptococci, the immune sera were able to promote phagocytosis and killing of only the homologous type 24 streptococci, indicating that the humoral responses to the synthetic peptide fragment are type-specific.

Agar gel diffusion tests using the immune rabbit sera gave precipitin arcs between the polylysine conjugates of the synthetic and the native CB7, as well as with the intact pepsin-extracted type 24 M protein (pep M24) molecule, which further confirms that the synthetic and native CB7 peptide fragments are immunochemically identical. Neither type 5 nor type 6 M protein was immunoreactive with anti-CB7 antiserum. None of the S-CB7 immune sera tested were reactive with frozen sections of human heart tissue assayed by immunofluorescence as previously described².

To demonstrate the protective capacity of the antisera against S-CB7, mice were passively immunized with a pool of the immune rabbit sera and challenged after 24 h with live type 24 or type 6 streptococci. The results (Table 1) clearly show the type-specific protective capacity of the immune sera and indicate that the S-CB7 peptide contains at least one protective antigenic determinant of type 24 streptococci.

To determine whether protective antigenic determinants resided in yet smaller peptide fragments, the CB7 peptide was cleaved at its arginine residue by trypsin, after blocking lysyl residues with recrystallized maleic anhydride (to a molar excess

of 20 over the total number of lysyl residues)^{4,7} After digestion with trypsin (TPCK (tosyl-phenylethyl-chloromethyl-ketone)-treated, Worthington) at an enzyme/substrate ratio of 1:50 (w/w) in 0.05 M NH_4HCO_3 , pH 8.3, the lysyl residues were demaleylated using pyridine/acetate (1:10), pH 3.0 at 60 °C for 6 h. The demaleylated peptides were then HPLC-separated on a column of Ultrasphere ODS2 equilibrated with 0.01 M phosphate buffer, pH 7.2, and eluted on a gradient of 0–40% acetonitrile⁸. In this way, a 12-residue COOH-terminal and a 23-residue NH_2 -terminal peptide were purified and then tested for their ability to inhibit opsonic antibodies⁴ (see Fig. 1 legend). For 50% inhibition of opsonization of type 24 streptococci, 9 nmol of the 23-residue, NH_2 -terminal peptide and 20 nmol of the 12-residue, COOH-terminal peptide of CB7 were needed, compared with only 1.6 nmol of a mixture of the two peptides and 0.8 nmol of the uncleaved CB7 peptide. The greater activity

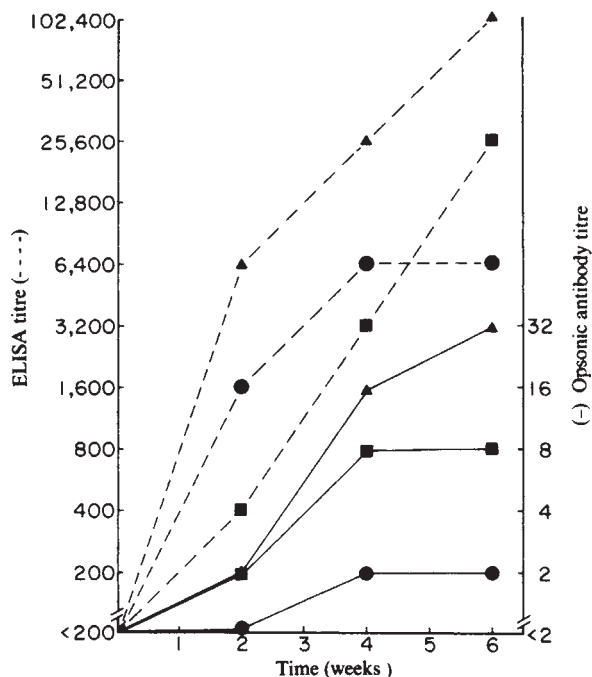


Fig. 1 Immune responses in three rabbits (▲, ■, ●) as measured by ELISAs (---) or opsonic antibody (—) assays after immunization with 24 nmol of S-CB7 conjugated to polylysine and emulsified in complete Freund's adjuvant (CFA) followed a week later by the same dose emulsified in CFA. Each symbol represents a different rabbit (see also Table 1 legend). Sera collected at 2-week intervals were assayed for antibodies to type 24 M protein adsorbed to the walls of plastic cuvettes by the ELISA method as previously described^{2,6}, except that semi-automatic methods were used (EIA PR50 automatic analyser and injector, Gilford). ELISA titres are expressed as the reciprocal of the highest dilution of serum giving an absorption >0.1 at 405 nm. Opsonic antibodies were assayed as described elsewhere¹. Briefly, the test mixture consisted of 0.4 ml of fresh heparinized (10 U ml^{-1}) human blood, 0.05 ml of a standard suspension of streptococci and 0.05 ml of various dilutions of test serum. The number of streptococcal units per leukocyte was ~10. The percentage of neutrophilic leukocytes counted that ingested one or more bacteria was estimated by microscopic examination of stained smears prepared from a drop of test mixture after incubation for 30 min. The opsonic antibody titres are expressed as the reciprocal of the highest twofold dilution of test serum in three separate tests that promoted phagocytosis of streptococci in $\geq 10\%$ of the neutrophils counted after incubation at 37 °C for 30 min; the same organisms in the presence of pre-immune rabbit serum were phagocytosed by $\leq 2\%$ of neutrophils in each test. Antisera giving titres >1:4 all produced phagocytosis in the range 40–70% when undiluted. The results of these phagocytosis tests were confirmed by indirect bactericidal tests performed as previously described¹⁸. Type specificity of the sera was confirmed by the failure of the S-CB7 immune sera to promote phagocytosis and killing of heterologous type 5 and type 6 streptococci.

Table 1 Protection of mice against challenge infections with type 24 streptococci by sera of rabbits immunized with S-CB7

Serum used to immunize mice passively	LD ₅₀ in mice challenged with:	
	Type 6 streptococci	Type 24 streptococci
Preimmune serum	<500 (2/15)	<500 (3/15)
Pooled (three rabbits) immune anti-S-CB7 serum	<500 (2/15)	3,500,000 (14/15)

Three rabbits were immunized intracutaneously with 25 nmol of S-CB7 conjugated to polylysine (MW ~35,000) and emulsified in CFA as previously described⁴. The initial immunizing dose was followed 1 week later by the same dose emulsified in incomplete Freund's adjuvant and injected subcutaneously. Preimmune and immune sera, obtained before immunization and 6 weeks after the initial immunizing dose, respectively, were pooled and white Swiss mice injected intraperitoneally with 0.2 ml of either serum. The mice were challenged 24 h later by the same route with various doses $500\text{--}4 \times 10^6$ colony-forming units of type 6 or type 24 streptococci. The survivals were recorded over a 7-day period and are shown in parentheses as the number of survivors per number of challenged mice.

of the mixture of the two peptide fragments of CB7 indicates that CB7 contains at least two distinct type-specific protective determinants, and that one of these resides in a peptide containing only 12 amino acid residues. The synthetic dodecapeptide overlapping these two peptides (residues 18–29) had no opsonic inhibitory effect in doses as high as 100 nmol, suggesting that neither immunodeterminant is included in this dodecapeptide. However, the possibility that the COOH-terminal triglycine residues may interfere with antibody binding at that end of the dodecapeptide has not been excluded.

In contrast to the type specificity of the humoral immune responses to S-CB7, the cellular immune responses were highly cross-reactive. The lymphocytes of S-CB7-immunized rabbits were equally responsive to a heterologous type 5 M protein and an homologous type 24 M protein (Table 2). Moreover, immunization of rabbits with the synthetic dodecapeptide (S18–29CB7), although not providing effective humoral immunity, induced cellular immunity to both serotypes of M protein similar to that seen after immunization with S-CB7 (Table 2). The lymphocytes from none of the animals responded to S-CB7 or S18–29CB7, indicating that these peptides were of insufficient molecular size to elicit the *in vitro* blastogenic response of sensitized lymphocytes⁹.

Recently Audibert *et al.*¹⁰ actively immunized laboratory animals against diphtheria toxin using a chemically synthesized oligopeptide. However, we believe that our results are the first to show that a synthetic polypeptide antigen can raise antibodies which promote phagocytosis and killing of a bacterial pathogen. The synthetic CB7 of type 24 streptococcal M protein represents a 35-amino acid fragment of the parent molecule, which consists of 376 amino acid residues¹. The molecule has previously been shown to consist of repeating covalent structures, the first 20 residues of CB7 being identical to the corresponding regions of four other peptide fragments (CB3, CB4, CB5 and CB6)¹. Two additional fragments (CB1 and CB2), each with a MW of ~10,000, have NH_2 -terminal amino acid sequences which are different from the smaller peptide fragments but are identical to each other and to the intact pepsin-extracted M protein (pep M24) molecule for at least the first 27 residues^{1,11}.

In addition, the 12 COOH-terminal residues of CB1 and CB2 were found to be identical to those of each of the smaller peptides (unpublished data). It was particularly intriguing that one of the protective determinants was shown to reside in this 12-residue peptide, which is repeated seven times in the pep M24 molecule. Because of the repetitive nature of the primary structure of the M protein molecule it is not surprising that antibody directed against such a small segment of the molecule proved to be opsonic and presumably protective. The repetition of the determinant would provide multiple sites of interaction

Table 2 Blastogenic responses of lymphocytes from rabbits immunized with S-CB7 or S18-29CB7 fragments of type 24 M protein

Rabbit	Immunizing antigen	Mean c.p.m. ($\pm$ s.e.m.) of lymphocytes cultured with:		
		Control	M5	M24
8026	S-CB7	98 $\pm$ 23	3,743 $\pm$ 411	4,721 $\pm$ 479
8027	S-CB7	181 $\pm$ 28	12,684 $\pm$ 1,454	829 $\pm$ 124
8028	S-CB7	156 $\pm$ 31	43,391 $\pm$ 3,094	36,449 $\pm$ 5,416
7919	S18-29CB7	415 $\pm$ 27	11,521 $\pm$ 3,081	12,038 $\pm$ 243
7921	S18-29CB7	135 $\pm$ 31	17,056 $\pm$ 3,081	3,575 $\pm$ 1,265
8020	S18-29CB7	705 $\pm$ 59	12,487 $\pm$ 918	4,916 $\pm$ 362
8007	CFA (control)	76 $\pm$ 102	64 $\pm$ 12	40 $\pm$ 2
8008	CFA (control)	31 $\pm$ 11	153 $\pm$ 25	83 $\pm$ 25

Heparinized (100 U ml⁻¹) peripheral blood was obtained from rabbits 2–6 weeks after immunization (see Table 1 legend) by cardiac puncture. Mononuclear cells were isolated by Ficoll-Hypaque gradient centrifugation as described elsewhere^{17,19}. Lymphocytes were washed three times and resuspended in RPMI 1640 (Gibco) supplemented with penicillin (100 U ml⁻¹), streptomycin (100 μ g ml⁻¹), L-glutamine (2 mM) and HEPES buffer (25 mM). Lymphocytes (2×10^5) were incubated at 37 °C with 50 μ g per culture of each antigen tested in 96-well microculture plates (Falcon Plastics) in a total volume of 200 μ l supplemented with 5% heat-inactivated fetal calf serum. Control cultures were incubated in the same volume of medium without antigen. Eighteen hours before collecting, 1 μ Ci of ³H-thymidine (specific activity 2 Ci mmol⁻¹, Research Products International) in 25 μ l culture medium was added to each well. All cultures were collected after 5 days using multiple automated sample harvester and the cells assayed for radioactivity in a liquid scintillation counter as described elsewhere^{17,20}. CFA, complete Freund's adjuvant. Control animals were injected with the same volume of CFA emulsified in 0.15 M NaCl without antigen.

with the immunoglobulin molecules. Such interaction at multiple sites seems to be necessary for optimal opsonization of M protein-containing streptococcal cells¹².

The significance of the cell-mediated cross-reactions between synthetic peptide fragments and heterologous M proteins is unclear; structural similarities between the M proteins^{11,13–16} may account for these cross-reactions as well as for the high degree of cellular immunity to various M proteins recently observed in lymphocytes from human adults as well as cord blood of newborn infants¹⁷.

The immunogenicity of small peptide fragments is encouraging for the development of safe and effective vaccines against those streptococcal infections that initiate rheumatic fever and rheumatic heart disease². The efficacy of very small peptides would permit disposal of a large portion of the M protein molecule and, therefore, should reduce the chances of eliciting immunological cross-reactions against host tissues. Thus, the continued identification of peptide structures responsible for protective immunity should yield a pool of small peptides that may eventually be synthesized and administered safely to humans as vaccine broadly protective against many serotypes of streptococci, particularly against those strains that trigger post-streptococcal sequelae.

We thank Edna Chiang, Ella Pinals, Elizabeth Ryan and Elizabeth Rowland for technical assistance and Johnnie Smith for secretarial assistance. This work was supported by funds from the US Veterans Administration and by USPHS grants AI-10085, AI-13550 and AM-16506. E.H.B. is the recipient of a Medical Investigatorship Award from the US Veterans Administration and W.A.S. is the recipient of a USPHS Young Investigatorship Award (DE05773-01).

Received 25 February; accepted 28 May 1981.

1. Beachey, E. H., Seyer, J. M. & Kang, A. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3163–3167 (1978).
2. Beachey, E. H., Stollerman, G. H., Johnson, R. H., Ofek, I. & Bisno, A. L. *J. exp. Med.* **150**, 862–877 (1979).
3. Stollerman, G. H. *Rheumatic Fever and Streptococcal Infection* (Grune & Stratton, New York, 1975).
4. Beachey, E. H., Seyer, J. M. & Kang, A. H. *J. biol. Chem.* **255**, 6284–6289 (1980).
5. Merrifield, R. B. *J. Am. chem. Soc.* **85**, 2149–2155 (1963).
6. Dale, J. B., Ofek, I. & Beachey, E. H. *J. exp. Med.* **151**, 1026–1038 (1980).
7. Butler, P. J. G., Harris, J. I., Hartley, B. S. & Leberman, R. *Biochem. J.* **112**, 679–689 (1969).
8. River, J. E. *J. Liq. Chromatogr.* **1**, 343–353 (1978).
9. Singh, B. *et al. J. Immun.* **124**, 1336–1342 (1980).
10. Audibert, F. *et al. Nature* **289**, 593–594 (1981).
11. Beachey, E. H. & Seyer, J. M. *Bacterial Vaccines* (eds Robbins, J. B. *et al.*) (Decker, New York, in the press).
12. Fischetti, V. A. *J. exp. Med.* **146**, 1108–1123 (1977).
13. Beachey, E. H., Seyer, J. M. & Kang, A. H. in *Streptococcal Diseases and the Immune Response* (eds Zabriskie, J. B. & Read, S. E.) 149–160 (Academic, New York, 1980).
14. Seyer, J. M., Kang, A. H. & Beachey, E. H. *Biochem. biophys. Res. Commun.* **92**, 546–553 (1980).
15. Manjula, B. N. & Fischetti, V. A. *J. Immun.* **124**, 261–267 (1980).
16. Manjula, B. N. & Fischetti, V. A. *J. exp. Med.* **151**, 695–707 (1980).

17. Dale, J. B., Simpson, W. A., Ofek, I. & Beachey, E. H. *J. Immun.* **126**, 1499–1505 (1980).
18. Beachey, E. H. & Stollerman, G. H. *J. exp. Med.* **134**, 351–365 (1971).
19. Boyum, A. *Scand. J. clin. Lab. Invest.* **21**, 77–89 (1968).
20. Beachey, E. H. *et al. J. Immun.* **122**, 189–195 (1979).

Binding of MLR precursors requires immunosorbents which present both self- and alloantigen

Heather Y. Schnagl & William Boyle

Department of Microbiology, University of Melbourne, Parkville, Victoria 3052, Australia

The response of T lymphocytes to antigens of the major histocompatibility complex (MHC) in a species is commonly measured by two assays. The mixed lymphocyte reaction (MLR) measures proliferating lymphocytes (T_p), and cell-mediated cytotoxicity (CMC) measures the cytotoxic T cells (T_c) generated, which can destroy tissue cells bearing the alloantigens against which the response is directed¹. T_p and T_c are thought to be derived from different precursors, the T_p-P and T_c-P respectively, and each seem to have different requirements for activation. Initially it was claimed that T_p-P and T_c-P were coded to respond to antigens specified by different genes in the MHC² but this distinction has since been shown not to be absolute^{3–6}. We previously demonstrated⁷ that T_p-P and T_c-P had different antigen-binding characteristics as T_c-P could be bound in a haplotype-specific manner to allogeneic monolayers but T_p-P were not bound in these conditions. We report here further experiments which show that these two populations have qualitatively distinct antigen recognition properties, by demonstrating that T_p-P will bind only when self- plus alloantigen are presented by a monolayer, whereas T_c-P will bind when alloantigen alone is present.

Our original observation⁷ that monolayers of preincubated spleen cells specifically absorbed cytotoxic precursors, T_c-P, but failed to bind precursors of proliferative cells, T_p-P, has been confirmed in many experiments. Because the proliferative response seen in MLR reactions had been shown to be strongest when stimulated by H-2I region antigens, it seemed possible that T_p-P did not bind because monolayers of spleen cells were used and not all these cells expressed Ia determinants. Poly-L-lysine monolayers of B-lymphocyte blasts were therefore prepared from B10 mice, then spleen cells from B10.BR mice incubated on these for 1 h. The nonadherent cells were recovered and cultured for MLR and CMC production against X-irradiated B10 and B10.D2 cells. Table 1 shows that the results were the same as those obtained with spleen cell monolayers; T_c-P were bound but not T_p-P. Similar experiments

Table 1 Absorption on B-cell blast monolayers and its effect on the proliferative and cytotoxic response of the nonadherent B10.BR responder population to stimulation with X-irradiated B10 spleen cells

Responder	Absorbing LPS blast monolayer	Stimulator	Proliferative	Cytotoxic titre	
			Response (c.p.m. ³ H-T)	H-2 ^b target	H-2 ^d target
B10.BR	—	X-B10	10,500±950	16.7	5.7
B10.BR	—	X-B10.D2	19,800±1,500	<1	23.8
B10.BR	—	X-B10.BR	480±60		
B10.BR	B10	X-B10	10,800±680	<1	1.7
B10.BR	B10	X-B10.D2	16,900±660	<1	23.3
B10.BR	B10	X-B10.BR	500±50		
X-B10.BR	B10	X-B10	230±20		
X-B10.BR	B10	X-B10.D2	130±10		

B10 spleen cells (3×10^6) were cultured in 75 cm² flasks (Falcon no. 3024) in a final volume of 75 ml of Eagle's minimum essential medium plus non-essential amino acids, supplemented with 10% v/v fetal calf serum (FCS) and 7×10^{-5} M 2-mercaptoethanol (MEM/FCS) containing $10 \mu\text{g ml}^{-1}$ of lipopolysaccharide (LPS, Difco no. 3122-25) at 37 °C with 10% CO₂ in air for 3 days to generate B-cell blasts. After washing, $1-2 \times 10^7$ blast cells were attached to poly-L-lysine-coated dishes as described previously⁷. B10.BR responder spleen cells, or in one case X-irradiated B10.BR (X-B10.BR) spleen cells as controls, were absorbed on the resultant monolayers as previously described⁷. Briefly, B10.BR responder spleen cells (2×10^7) in 2 ml MEM/FCS were added to each monolayer. The dishes were centrifuged at 100g for 5 min then incubated at 37 °C with 10% CO₂ in air for 1 h. The nonadherent B10.BR responder cells were recovered by brief agitation, washed once, and 2×10^7 were resuspended in 8 ml MEM/FCS and divided equally between two 25 cm² flasks (Falcon no. 3013). Control cultures containing 10^7 unfractionated B10.BR spleen cells in 4 ml MEM/FCS were similarly prepared. To each flask were added 10^7 X-irradiated (1,000 rad) stimulator spleen cells in 1 ml. Stimulator cells of the monolayer strain (X-B10, H-2^b) were added to one of each pair of flasks representing the nonadherent population from a single dish, while stimulators of an unrelated strain (X-B10.D2, H-2^d) or of the responder strain (X-B10.BR) were added to the remaining flask. All cultures were prepared in duplicate and incubated upright at 37 °C with 10% CO₂ in air for 5 days. To determine the proliferative response, duplicate 0.3-ml samples of each culture after 3 days incubation were transferred to a microtitre tray (Sterilin no. M29 ARTL). After centrifugation at 100g for 5 min, half the supernatant was discarded and 25 μl of medium containing 0.5 μCi of ^3H -thymidine ($^3\text{H-T}$) were added to each well. After incubation in a humidified box with 10% CO₂ in air for 18 h, cultures were transferred to microtubes and treated according to the method of Smart *et al.*¹⁶. Results obtained were expressed as c.p.m. $^3\text{H-T}$ and represent the mean $\pm$ s.d. of quadruplicate determinations. Cytotoxicity was determined after 5 days incubation using a microassay described elsewhere¹⁷. Duplicate cultures were pooled, washed and resuspended in 0.3 ml RPMI 1640 supplemented with 10% v/v FCS (RPMI/FCS). All T_c populations were titrated against ^{51}Cr -labelled tumour target cells of both stimulator haplotypes. The results obtained were expressed as % lysis $\pm$ s.d. of triplicate determinations. Per cent specific lysis represents the % lysis minus the spontaneous release of label by target cells incubated with medium alone. The cytotoxic titre for a given T_c population was calculated as the reciprocal of the highest dilution of the T_c population which gave 15% specific lysis of the target concerned.

using mouse embryonic fibroblasts which have also been shown to express I-region antigens^{8,9} also failed to show depletion of T_p-P.

Although T-cell responses to MHC alloantigens have always been considered not to require self-antigen presentation, in the absence of any other effective system for binding T_p-P, we examined the effect of using absorbent monolayers which co-presented both allo- and self-antigen. Table 2 shows the results of absorbing B10.BR spleen cells separately on to monolayers composed of preincubated splenic lymphocytes of B10.BR, BALB/c or (B10.BR $\times$ BALB/c)F₁ hybrid origin. Neither the B10.BR nor BALB/c monolayers bound T_p-P, as shown by the unimpaired MLR response of the nonadherent B10.BR cells when stimulated with X-irradiated BALB/c or B10 cells. The BALB/c monolayers were adequate as shown by the fact that, although T_p-P were not bound, T_c-P directed against BALB/c were bound (data not shown). However, B10.BR cells which had been exposed to the (B10.BR $\times$ BALB/c)F₁ monolayer were found to be depleted of both T_c-P and T_p-P when stimulated by BALB/c alloantigens. The T_p-P depletion was apparently specific as the T_p response to B10 alloantigens was unimpaired.

To ensure that the reduction in ability to respond to BALB/c antigens by the B10.BR cells, which were nonadherent to the (B10.BR $\times$ BALB/c)F₁ monolayer, was due to binding of T_p-P and not to induction of a suppressor effect, non-F₁ monolayer-

adherent B10.BR cells were added to normal B10.BR spleen cells and stimulated with X-irradiated BALB/c cells. As can be seen in Table 3, no evidence was found for suppression by the B10.BR cells that were nonadherent to the F₁ monolayer. An alternative explanation of the results is that T_p-P were either bound or inactivated nonspecifically by exposure to the F₁ monolayer and that the apparent specificity of the undiminished anti-B10 response was attributable to F₁ monolayer cells which were mounting an anti-B10 response. However, this was not the case as the cells which incorporated $^3\text{H-T}$ on stimulation by B10 alloantigens were susceptible to complement-mediated lysis with anti-B10.BR serum but not with anti-BALB/c serum (Table 4). It was therefore concluded that the reduction in proliferation was due to binding of the T_p-P to the F₁ monolayer.

The specificity of T_p-P depletion was confirmed and investigated further by exposing B10.BR spleen cells to monolayers of B10.BR, BALB/c, B10, (B10.BR $\times$ BALB/c)F₁, (B10.BR $\times$ B10)F₁ or (B10 $\times$ BALB/c)F₁ spleen cells. The results are summarized in Table 5. Absorption of anti-BALB/c T_p-P occurred only on (B10.BR $\times$ BALB/c)F₁ monolayers. The depletion is specific for the anti-BALB/c response as the nonadherent B10.BR cells gave an undiminished MLR response to B10 alloantigen stimulation. The specificity of the depletion technique is further illustrated by the demonstration that absorption on a (B10.BR $\times$ B10)F₁ monolayer depleted the anti-B10 proliferative response but not the anti-BALB/c response. This experiment also defines further the self requirement, because B10.BR were not depleted of anti-BALB/c T_p-P when exposed to a monolayer of (B10 $\times$ BALB/c)F₁, indicating that the self-antigen recognized must be coded in the H-2 region. This has been confirmed by further experiments (data not shown) in which B10 responder cells

Table 2 Absorption of B10.BR responders on preincubated (B10.BR $\times$ BALB/c)F₁ monolayers specifically depletes the proliferative response of the nonadherent responder population to stimulation with X-irradiated BALB/c spleen cells

Responder	Absorbing monolayer	Stimulator	Proliferative response (c.p.m. $^3\text{H-T}$)
B10.BR	—	X-BALB/c	6,560 ± 480
B10.BR	—	X-B10	5,310 ± 510
B10.BR	—	X-B10.BR	500 ± 80
B10.BR	B10.BR	X-BALB/c	6,710 ± 310
B10.BR	B10.BR	X-B10	5,450 ± 470
B10.BR	B10.BR	X-B10.BR	510 ± 50
B10.BR	BALB/c	X-BALB/c	7,890 ± 960
B10.BR	BALB/c	X-B10	6,420 ± 400
B10.BR	BALB/c	X-B10.BR	510 ± 60
B10.BR	(B10.BR $\times$ BALB/c)F ₁	X-BALB/c	310 ± 120
B10.BR	(B10.BR $\times$ BALB/c)F ₁	X-B10	5,850 ± 610
B10.BR	(B10.BR $\times$ BALB/c)F ₁	X-B10.BR	390 ± 40

B10.BR responder spleen cells were absorbed onto B10.BR, BALB/c (H-2^d) or (B10.BR $\times$ BALB/c)F₁ spleen cell monolayers (preincubated for 4 h) as described in Table 1. The resulting nonadherent responder cells were stimulated with X-irradiated BALB/c (X-BALB/c), B10 (X-B10) or B10.BR (X-B10.BR) spleen cells. After 3 days incubation the proliferative response was determined as described in Table 1 legend.

Table 3 Inability of B10.BR responders nonadherent to (B10.BR $\times$ BALB/c)F₁ preincubated monolayers to suppress the proliferative response of normal, unfractionated B10.BR responders to stimulation with X-irradiated BALB/c

Responder population			Proliferative response (c.p.m. $^3\text{H-T}$)
B10.BR	X-B10.BR	B10.BR nonadherent to F ₁ * monolayer	
5×10^6	—	5×10^6	2,930 ± 250
5×10^6	5×10^6	—	2,780 ± 110
—	5×10^6	5×10^6	230 ± 70

B10.BR responder cells nonadherent to a (B10.BR $\times$ BALB/c)F₁ monolayer (5×10^6) were mixed with an equal number of unfractionated B10.BR spleen cells. These were stimulated with 10^7 X-irradiated BALB/c spleen cells as were control cultures containing 5×10^6 unfractionated B10.BR responder spleen cells and 5×10^6 X-irradiated B10.BR spleen cells.

* F₁ = (B10.BR $\times$ BALB/c)F₁.

could be depleted of anti-BALB/c T_p-P by exposure to a (B10×BALB/c)F₁ spleen monolayer but not a (B10.BR×BALB/c)F₁ monolayer.

The specificity of binding of T_p-P by monolayers presenting self- plus alloantigen has been tested using a number of parental and F₁ strains. These have confirmed the general observation that T_p-P were bound only when the responder and absorbing F₁ monolayer shared a common haplotype, and the depletion was specific for the haplotype of the other parent. In addition, we have found that monolayers composed of mixtures of responder and stimulator spleen cells would not absorb anti-stimulator T_p-P. The level of depletion achieved by absorption on semi-allogeneic F₁ monolayers has varied for different combinations of responder-monolayer mouse strains but has been consistent for any single combination. This is not surprising as we found that even for T_c-P binding, the level of depletion depends on the time of preincubation before the monolayer is constructed and this differs from strain to strain. All the experiments reported here using F₁ cells have involved only a 4-h preincubation which may not be optimal for all combinations.

Table 4 Proliferative response of the B10.BR nonadherent population to stimulation by X-irradiated B10 is not due to detached (B10.BR×BALB/c)F₁ monolayer cells

Responder	Preincubated monolayer	Stimulator	Treatment	Proliferative response (c.p.m. ³ H-T)
B10.BR	—	X-B10	C' only	4,980±220
B10.BR	—	X-B10	Anti-BALB/c+C'	4,560±190
B10.BR	—	X-B10	Anti-B10.BR+C'	2,280±90
F ₁ *	—	X-B10	C' only	6,680±290
F ₁	—	X-B10	Anti-BALB/c+C'	3,050±120
F ₁	—	X-B10	Anti-B10.BR+C'	3,100±110
B10.BR	F ₁	X-B10	C' only	4,380±210
B10.BR	F ₁	X-B10	Anti-BALB/c+C'	4,550±180
B10.BR	F ₁	X-B10	Anti-B10.BR+C'	1,740±80

B10.BR, (B10.BR×BALB/c)F₁ or B10.BR responders that were nonadherent to preincubated (B10.BR×BALB/c)F₁ monolayers were stimulated with X-irradiated B10 (X-B10) spleen cells. Samples of 3-day cultures were incubated with ³H-T as described in Table 1 legend. After labelling, replicate samples were pooled, washed and resuspended in 0.2 ml RPMI/FCS. To three 50-μl aliquots the following were added: (1) 50 μl of RPMI/FCS; (2) 50 μl of a 1:4 dilution of B10.BR anti-BALB/c serum; (3) 50 μl of a 1:4 dilution of a BALB/c anti-B10.BR serum. After incubation for 30 min the tubes were centrifuged and supernatants discarded; 0.2 ml of a 1:8 dilution of rabbit complement (C') was added to each tube and incubated for a further 30 min. After washing twice in RPMI/FCS the samples were collected and processed as described in Table 1 legend.

* F₁ = (B10.BR×BALB/c)F₁.

Table 5 Specific depletion of T_p-P requires the responder and monolayer strains to share an H-2 haplotype

Responder	Absorbing monolayer	Stimulator	Proliferative response (c.p.m. minus control values)
B10.BR	—	X-BALB/c	11,020±710
B10.BR	—	X-B10	20,660±860
B10.BR	B10.BR	X-BALB/c	10,770±810
B10.BR	B10.BR	X-B10	19,770±1,020
B10.BR	BALB/c	X-BALB/c	10,950±720
B10.BR	BALB/c	X-B10	19,750±690
B10.BR	B10	X-BALB/c	11,080±570
B10.BR	B10	X-B10	18,020±790
B10.BR	(B10.BR×BALB/c)F ₁	X-BALB/c	80±70
B10.BR	(B10.BR×BALB/c)F ₁	X-B10	19,340±900
B10.BR	(B10.BR×B10)F ₁	X-BALB/c	10,480±330
B10.BR	(B10.BR×B10)F ₁	X-B10	900±110
B10.BR	(B10×BALB/c)F ₁	X-BALB/c	10,370±520
B10.BR	(B10×BALB/c)F ₁	X-B10	20,640±1,020

B10.BR responder spleen cells were absorbed, as described in Table 1 legend, on monolayers of B10.BR, BALB/c, B10, (B10.BR×BALB/c)F₁, (B10.BR×B10)F₁, or (B10×BALB/c)F₁ spleen cell monolayers that had been preincubated for 4 h. The resultant nonadherent responder cells were stimulated with X-irradiated BALB/c, B10 or B10.BR spleen cells. After 3 days incubation the proliferative response was determined as described in Table 1 legend. The results are expressed as c.p.m. minus control values (stimulated with X-B10.BR).

The importance of the experiments reported here is that, in addition to clarifying the distinction between antigen recognition by T_p-P and T_c-P, they present the first evidence of the need for self-antigen recognition by T cells coded to respond to alloantigens of the MHC. It is already well established that T cells responsive to non-MHC antigens require the extraneous antigen to be presented together with self-antigens¹⁰⁻¹⁵. This distinction has been the basis of suggestions that alloreactive cells might, in fact, include multiple clones coded to self plus various extraneous antigens. This view will require some revision if the need for self- plus alloantigen demonstrated here is found to be generally applicable to the activation of cells which proliferate in response to MHC alloantigens. It should be stressed that in our experiments we have measured binding of precursors, not activation, but preliminary tests suggest that the same features apply to activation.

This work was supported by an Australian NH and MRC grant to W.B. H.Y.S. was supported by a Commonwealth Postgraduate Research Award. We thank Julie Ogilvie for technical assistance and the animal house staff for the provision of the mouse strains.

Received 13 October 1980; accepted 13 May 1981.

- Wagner, H., Rölinghoff, M. & Nossal, G. J. V. *Transplant Rev.* **17**, 3-36 (1973).
- Bach, F. H. *et al. Science* **180**, 403-406 (1973).
- Widmer, M. B., Alter, B. J., Bach, F. H. & Bach, M. L. *Nature new Biol.* **242**, 239-241 (1973).
- Klein, J., Hauptfeld, M. & Hauptfeld, V. *Immunogenetics* **1**, 45-56 (1974).
- Nabholz, M. *et al. Eur. J. Immun.* **5**, 594-599 (1975).
- Wagner, H., Götz, D., Ptschelinzew, L. & Rölinghoff, M. J. *exp. Med.* **142**, 1477-1487 (1975).
- Schnagl, H. Y. & Boyle, W. *Nature* **279**, 331-333 (1979).
- Hämmerling, G. J., Mauve, G., Goldberg, H. & McDevitt, H. O. *Immunogenetics* **1**, 428-437 (1975).
- McDevitt, H. O., Delovitch, T. L., Press, J. L. & Murphy, D. B. *Transplant Rev.* **30**, 197-235 (1976).
- Katz, D. H., Hamaoka, T. & Benacerraf, B. J. *exp. Med.* **137**, 1405-1418 (1973).
- Zinkernagel, R. M. & Doherty, P. C. *Nature* **248**, 701-702 (1974).
- Shearer, G. M., Rhen, T. C. & Garbarino, C. A. J. *exp. Med.* **141**, 1348-1364 (1975).
- Gordon, R. D., Simpson, E. & Samelson, L. E. J. *exp. Med.* **142**, 1108-1120 (1975).
- Bevan, M. J. *exp. Med.* **142**, 1349-1364 (1975).
- Miller, J. F. A. P., Vadas, M. A., Whitelaw, A. & Gamble, J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2486-2490 (1976).
- Smart, I. J., Boyle, W. & Morris, P. J. *Tissue Antigens* **11**, 20-29 (1978).
- Schnagl, H. Y. & Boyle, W. J. *Immun. Meth.* **24**, 79-88 (1978).

Adrenaline activation of phosphofructokinase in rat heart mediated by α-receptor mechanism independent of cyclic AMP

M. G. Clark & G. S. Patten

Commonwealth Scientific and Industrial Research Organization, Division of Human Nutrition, Kintore Avenue, Adelaide, South Australia 5000, Australia

We recently reported¹ that adrenaline treatment of the perfused rat heart produced a marked activation of phosphofructokinase when assayed at low substrate concentration using a new isotopic assay². As muscle phosphofructokinase has been reported to be phosphorylated by the cyclic AMP-dependent protein kinase³, activation of phosphofructokinase might be expected to result from phosphorylation mediated by the β-receptor/adenylyl cyclase/cyclic AMP-dependent protein kinase sequence. We now report that this is not so, but that activation of phosphofructokinase in the perfused rat heart occurs independently of phosphorylation, by a cyclic AMP-independent mechanism, and that this process seems to be mediated predominantly through the α-adrenergic receptors. We believe this represents the first example showing that not only in liver^{4,5}, but also in a muscular tissue, occupancy of an α-adrenergic receptor changes the activity of a glycolytic enzyme. In addition, the present findings imply a coordination of muscle glycolysis with hepatic glucose output.

Figure 1 shows the effect of varying ATP concentration on extract phosphofructokinase activity from control and adrenaline-treated hearts. The adrenaline-activated form of the enzyme as expressed in crude heart extracts displays higher activity than the control form at all ATP concentrations. This is most pronounced at ATP concentrations >0.5 mM, where the control form is almost totally inhibited. We have subsequently used this difference in sensitivity to inhibition by ATP to define an 'activity ratio' for phosphofructokinase where the control (unactivated) form of the enzyme is 0.2 ± 0.1 and the activated form of the enzyme is 0.6 ± 0.1 , representing the activity of phosphofructokinase determined at 1 mM ATP divided by that at 0.1 mM ATP (both at 0.01 mM hexose 6-phosphate).

Table 1 shows the effect of 5-min perfusions with adrenergic agonists or antagonists on the activity ratios of phosphofructokinase and glycogen phosphorylase in heart. For phosphofructokinase the basal activity ratio of 0.21 was increased about threefold by both adrenaline (10^{-5} M) and 10^{-5} M phenylephrine and remained elevated following gel filtration of the heart extracts (results not shown). Neither 10^{-5} M isoprenaline nor 10^{-5} M phenoxybenzamine had any effect, but 10^{-5} M propranolol increased the ratio to 0.35. As the increase produced by adrenaline was blocked by phenoxybenzamine but not by propranolol, it was concluded that activation of phosphofructokinase was mediated predominantly through the α -receptor. This view was further supported by the observation that the marked increase in phosphofructokinase activity ratio produced by 10^{-5} M phenylephrine was also blocked by 10^{-5} M phenoxybenzamine.

In contrast to a predominantly α -receptor-mediated activation of phosphofructokinase, the activation of phosphorylase was mediated by a β -receptor mechanism. Adrenaline (10^{-5} M and 5×10^{-7} M) and 10^{-5} M isoprenaline each increased the phosphorylase activity ratio ($a/a+b$) approximately fourfold. 10^{-5} M Phenylephrine had a marginal effect. Neither 10^{-5} M propranolol nor 10^{-5} M phenoxybenzamine had any stimulatory effect. The increases produced by adrenaline and isoprenaline were totally blocked by the β -antagonist, propranolol (10^{-5} M), but the adrenaline-mediated increase in phosphorylase activity ratio remained unaffected by the α -antagonist, phenoxybenzamine (10^{-5} M).

The marginal increase in phosphorylase activity ratio produced by the highest concentration of phenylephrine (10^{-5} M) was blocked by 10^{-5} M phenoxybenzamine. However, lower concentrations of phenylephrine ($\leq 10^{-6}$ M) significantly increased the activity ratio of phosphofructokinase but did not increase that of phosphorylase (Table 1).

Figure 2 shows the time course effects of 10^{-5} M isoprenaline, 10^{-5} M phenylephrine and 10^{-5} M naphazoline on the activity ratios of phosphofructokinase and phosphorylase as well as the

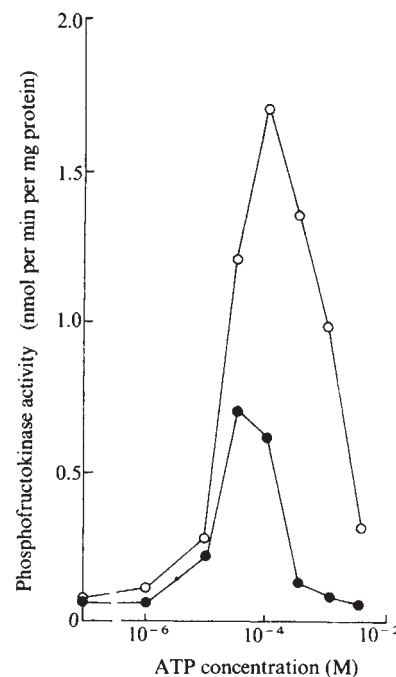


Fig. 1 Effect of ATP concentration on phosphofructokinase activity in extracts from control (●) and adrenaline-treated (○) hearts. Hearts of adult male hooded Wistar rats (200–250 g) that had been fed *ad libitum* on a standard diet were perfused in the Langendorff manner¹¹. The perfusion medium contained 118 mM NaCl, 4.8 mM KCl, 1.2 mM MgSO_4 , 1.2 mM phosphate, 25 mM NaHCO_3 , 3.0 mM CaCl_2 and 0.5 mM Na_2EDTA and was equilibrated with $\text{O}_2 + \text{CO}_2$ (95 + 5) to give a pH and $p\text{O}_2$ of 7.4 and 450 mm Hg, respectively. After a 10-min non-recirculating perfusion (pressure 60 mm Hg, temperature 37°C) the hearts were perfused in a recirculating manner with identical medium containing hormone or saline. After the appropriate time of perfusion (in this case 5 min) the hearts were freeze-clamped using tongs that had been precooled in liquid N_2 . The powdered hearts were kept at -70°C . Cell-free extracts of frozen hearts were prepared by homogenizing 1 part by weight of powdered heart with 10 vol of homogenizing buffer (100 mM Tris-HCl, pH 7.4, which contained 2 mM dithiothreitol) using an Ultra-Turrax homogenizer (speed 20,000 r.p.m. for 30 s). Consistent results for phosphofructokinase activity were obtained when frozen heart powder was added to preweighed tubes containing homogenizing buffer that had been frozen in liquid N_2 and the combined mixture homogenized during thawing at 5°C. Thawing of the powdered heart before homogenization resulted in abnormally high phosphofructokinase activities, presumably due to the action of endogenous activators. The homogenate was centrifuged for 30 s at 8,000g and kept on ice. In these conditions the activity increased by $\sim 10\%$ per hour. 'Control' heart before homogenization resulted in abnormally high phosphofructokinase activities, presumably due to the action of endogenous biological fructose 6-phosphate concentration using a newly developed radioisotopic procedure². The assay has been previously shown to be specific for phosphofructokinase and to give values for maximal catalytic activity identical to those obtained using the conventional spectrophotometric assay of 16 nmol per min per mg protein (ref. 2). The assays were conducted in disposable plastic tubes (70×10 mm) at 30°C and contained 10 μM glucose 6-phosphate, 3.3×10^4 d.p.m. [$5\text{-}^3\text{H}$]glucose 6-phosphate (prepared from [$5\text{-}^3\text{H}$]glucose²), ATP as indicated, 10 mM MgCl_2 and 20 μl of cell-free supernatant (0.075–0.1 mg of protein) in a total volume of 0.2 ml of homogenizing buffer. Reactions were started by the addition of the extract and terminated by the addition of 0.5 ml of Dowex 1 $\times$ 8 (200–400 mesh) borate resin (1:1 in water), vortexed, then centrifuged for 3 min at 1,300g. A 0.25-ml portion of the supernatant (containing only $^3\text{H}_2\text{O}$) was mixed with 2 ml of a Triton X-based water-accepting scintillant and counted for radioactivity. The activity of phosphofructokinase is expressed in nmol of [$5\text{-}^3\text{H}$]hexose 6-phosphate converted to $^3\text{H}_2\text{O}$ per min per mg of protein, and was calculated from the d.p.m. in $^3\text{H}_2\text{O}$ and the specific radioactivity of hexose 6-phosphate. The specific radioactivity of hexose 6-phosphate was calculated from the d.p.m. added as [$5\text{-}^3\text{H}$]glucose 6-phosphate and the sum of unlabelled glucose 6-phosphate and fructose 6-phosphate (added plus endogenous). Glucose 6-phosphate and fructose 6-phosphate were determined on neutralized perchlorate extracts of heart¹². The total concentration of hexose 6-phosphate in the assays was 10.7×10^{-6} and 14×10^{-6} M for control and adrenaline-treated extracts, respectively. Glycogen phosphorylase was measured using the filter disk assay of Gilboe *et al.*¹³ as modified by Birnbaum and Fain¹⁴ and was determined on extracts that had been passed through columns (0.7×2.5 cm) of Sephadex G-25. Protein was determined using the procedure of Lowry *et al.*¹⁵ with bovine serum albumin as the standard. A total of four hearts (two control and two adrenaline-treated) were used. Mean values are shown.

Table 1 The effect of adrenergic agonists and antagonists on the activity ratios of phosphofructokinase and glycogen phosphorylase

Additions	Phosphofructokinase	Phosphorylase
None	0.21 ± 0.01 (4)	0.16 ± 0.004 (4)
10^{-5} M adrenaline	0.60 ± 0.10 (14)	0.68 ± 0.027 (10)
5×10^{-7} M adrenaline	0.42 ± 0.02 (3)	0.60
10^{-5} M adrenaline	0.57 ± 0.05 (3)	0.14 ± 0.031 (3)
+ 10^{-5} M propranolol		
10^{-5} M adrenaline	0.43	0.66
+ 10^{-5} M phenoxybenzamine		
5×10^{-7} M adrenaline	0.23 ± 0.02 (3)	0.58
+ 10^{-5} M phenoxybenzamine		
10^{-5} M propranolol	0.35 ± 0.10 (3)	0.16 ± 0.004 (3)
10^{-5} M phenoxybenzamine	0.18 ± 0.06 (3)	0.15 ± 0.015 (3)
10^{-5} M isoprenaline	0.24 ± 0.04 (3)	0.72 ± 0.036 (3)
10^{-5} M isoprenaline	0.34 ± 0.02 (3)	0.15 ± 0.007 (3)
+ 10^{-5} M propranolol		
10^{-5} M phenylephrine	0.62 ± 0.01 (3)	0.31 ± 0.018 (3)
10^{-6} M phenylephrine	0.40 ± 0.02 (3)	0.17 ± 0.005 (3)
10^{-7} M phenylephrine	0.37 ± 0.09 (3)	0.16 ± 0.017 (3)
10^{-5} M phenylephrine	0.35 ± 0.03 (3)	0.16 ± 0.016 (3)
+ 10^{-5} M phenoxybenzamine		

The values are means $\pm$ s.e.m. with the number of hearts in parentheses.

tissue content of cyclic AMP. 10^{-5} M isoprenaline had no significant effect on the activity ratio of phosphofructokinase over the 15-min perfusion period. However, within 30 s there were marked increases in both cyclic AMP concentration and the activity ratio of phosphorylase. Both remained close to the maximal value throughout the 15-min non-recirculating perfusion. Phenylephrine (10^{-5} M), generally considered as predominantly an α -agonist but with some β activity in the rat^{5,6}, increased the activity ratio of phosphofructokinase from 0.2 at 0 min to a maximal value of 0.65 at 1.5 min. The activation was similar to the time course previously reported for adrenaline¹. Phenylephrine also produced an increase in the activity ratio of phosphorylase (100%). Naphazoline (10^{-5} M), an α -agonist⁷, increased the activity ratio of phosphofructokinase in a similar time-dependent manner to phenylephrine, reaching a maximum value of 0.55 at 1.5 min. Naphazoline had no effect on phosphorylase or cyclic AMP concentration.

The present findings suggest that short-term regulation of phosphofructokinase in rat heart by catecholamines is mediated by α -adrenergic receptors. This process may operate in anticipation of increased glucose released from the liver⁸ when catecholamines are released into the blood in emergency situations. The β -receptor component in the heart would lead to increased glycogenolysis, contractile rate and force. The α -receptor component would separately permit increased flux through glycolysis from either myocardial glycogen stores or extra-myocardial glucose. Such a mechanism for α -receptor-mediated, cyclic AMP-independent control of phosphofructokinase would complement the observation by Keely *et al.*⁶ of an α -mediated glucose transport system in rat heart.

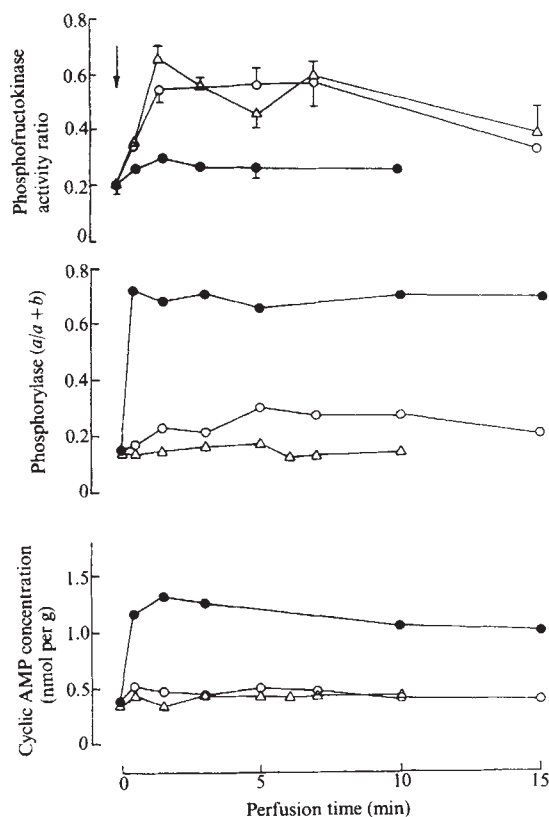


Fig. 2 Time course for the effects of α - and β -agonists on perfused rat heart. Hearts were perfused as described in Fig. 1. At the time indicated (arrow), medium containing 10^{-5} M isoprenaline ($\bullet$), 10^{-5} M naphazoline (Δ) or 10^{-5} M phenylephrine ($\circ$) was introduced. Non recirculating perfusion was continued until the hearts were freeze-clamped. The activity ratios of phosphofructokinase and glycogen phosphorylase were determined as described in Fig. 1 and the text. Solutions of agonists were freshly prepared in isotonic saline and mixed with warm (37°C) oxygenated buffer immediately before introduction into the perfusion apparatus. Cyclic AMP was determined¹⁶ on neutralized perchlorate extracts of the frozen heart powders. Where indicated mean values $\pm$ s.e.m. have been calculated (bars) from at least three hearts. Other values are the mean of duplicate determinations from single hearts.

Activation of phosphofructokinase in rat heart by α -agonists seems to involve mechanisms quite different from those controlling glucose output in rat liver. Exton and co-workers⁸ have provided strong evidence that α -receptor stimulation in liver leads to Ca^{2+} release from intracellular stores. This in turn is believed to activate the Ca^{2+} -dependent phosphorylase kinase that phosphorylates and activates phosphorylase. The present study indicates that activation of phosphorylase and phosphofructokinase are separate, suggesting a unique mechanism for phosphofructokinase activation. Furthermore, any increase in intracellular Ca^{2+} concentration should enhance contractility. In the present study, the α -agonist naphazoline, which activated phosphofructokinase, had no visible effect on either heart rate or force.

The mechanism by which α -agonists activate heart muscle phosphofructokinase is unknown. It seems unlikely that changes in activity can be related to changes in the concentration of intracellular effectors¹ even though a new activator of the liver enzyme has recently been described^{9,10}. Covalent modification including phosphorylation³ requires serious consideration as the change in activity is stable to gel filtration¹ or dialysis (unpublished observations), providing buffers contain EDTA and sulphhydryl reagents.

Received 24 March; accepted 20 May 1981.

- Clark, M. G. & Patten, G. S. *J. biol. Chem.* **256**, 27–30 (1981).
- Clark, M. G. & Patten, G. S. *J. molec. cell. Cardiol.* **12**, 1053–1064 (1980).
- Riquelme, P. T., Hosey, M. M., Marcus, F. & Kemp, R. G. *Biochem. biophys. Res. Commun.* **85**, 1480–1487 (1978).
- Feliú, J. E., Hue, L. & Hers, H. G. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2762–2766 (1976).
- Kemp, B. E. & Clark, M. G. *J. biol. Chem.* **253**, 5147–5154 (1978).
- Keely, S. L., Corbin, J. D. & Lincoln, T. *Molec. Pharmacol.* **13**, 965–975 (1977).
- Innes, I. R. & Nickerson, M. in *The Pharmacological Basis of Therapeutics* 4th edn (eds Goodman, L. S. & Gilman, A.) 478–523 (Macmillan, New York, 1970).
- Exton, J. H. *Biochem. Pharmacol.* **28**, 2237–2240 (1979).
- Van Schaftingen, E., Hue, L. & Hers, H. G. *Biochem. J.* **192**, 897–901 (1980).
- Claus, T. H., Schlumpf, J., Pilgis, J., Johnson, R. A. & Pilgis, S. J. *Biochem. biophys. Res. Commun.* **98**, 359–366 (1981).
- Neely, J. R., Liebermeister, H., Battersby, E. J. & Morgan, H. E. *Am. J. Physiol.* **36**, 413–459 (1974).
- Hohorst, H. J. in *Methods of Enzymatic Analysis* (ed. Bergmeyer, H. U.) 134–138 (Academic, New York, 1965).
- Gilboe, D. P., Larson, K. L. & Nuttall, F. Q. *Analyt. Biochem.* **47**, 20–27 (1972).
- Birnbaum, M. J. & Fain, J. N. *J. biol. Chem.* **252**, 528–535 (1977).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
- Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **67**, 305–312 (1970).

Evidence that LY-141865 specifically stimulates the D-2 dopamine receptor

K. Tsuruta, E. A. Frey, C. W. Grewe, T. E. Cote, R. L. Eskay* & J. W. Kebabian†

Experimental Therapeutics Branch, NINCDS, and *Laboratory of Clinical Science, NIMH, National Institutes of Health, Bethesda, Maryland 20205, USA

Available evidence suggests the existence of multiple categories of dopamine receptor^{1–4}. According to the classification scheme used here⁵, dopamine receptors can be divided into two general categories, D-1 and D-2. The dopamine receptors in the bovine parathyroid gland^{6–8} and carp retina^{9,10} exemplify the former category, and the dopamine receptors on the mammotrophs^{11,12} and melanotrophs^{13,14} of the pituitary gland exemplify the latter. Stimulation of a D-1 receptor enhances the formation of cyclic AMP, whereas stimulation of a D-2 receptor does not^{4,5}. The availability of dopaminergic agonists which specifically activate either a D-1 or a D-2 receptor would facilitate pharmacological investigation of the physiological responses to dopamine; however, to date, an agonist specific for the D-2 receptor has not

† To whom correspondence should be addressed.

been available. Dopamine and other catecholamines stimulate both types of receptor⁶⁻¹⁴; ergolines (such as lergotril and lisuride) stimulate the D-2 receptor¹³⁻¹⁵ but antagonize the D-1 receptor^{8,16-18}; some agonists acting on the D-2 receptor (such as apomorphine) both stimulate and block the D-1 receptor^{6,8}. Although one report claims that S-584 may selectively stimulate the dopamine receptor on mammothrophs¹⁹, according to another, S-584 stimulates the striatal dopamine-sensitive adenylate cyclase²⁰. Several recently synthesized compounds containing various portions of the ergoline ring system stimulate the D-2 receptor on mammothrophs²¹. We report here that one of these ergoline analogues, LY-141865 or '38, R = pro'²¹, stimulates the D-2 receptor in the intermediate lobe (IL) of the rat pituitary gland but does not interact with the D-1 receptor in carp retina.

The melanotrophs of the IL synthesize and secrete α -melanocyte-stimulating hormone (α -MSH). Stimulation of the dopamine receptor in the rat IL inhibits basal release of α -MSH^{13,14}. LY-141865 mimics the ability of dopamine and dopaminergic agonists to diminish the basal release of α -MSH from the IL. Thus, during a 3-h incubation of dispersed cells derived from the neurointermediate lobe, LY-141865 decreased, in a dose-dependent manner, the basal release of α -MSH (Fig. 1a). Half-maximal inhibition was achieved with ~ 30 nM LY-141865, maximal inhibition with a concentration ≥ 1 μ M.

The melanotrophs of the rat IL also possess a β -adreno-receptor, stimulation of which enhances adenylate cyclase activity¹⁵. Stimulation of the dopamine receptor diminishes basal adenylate cyclase activity and inhibits in a non-competitive manner the (-)isoprenaline-stimulated adenylate cyclase activity^{13-15,22}. Like other dopaminergic agonists, LY-141865 diminished the basal adenylate cyclase activity (Fig. 2, inset A) and the (-) isoprenaline-stimulated enzyme activity (Fig. 2 left, and inset B). Half-maximal inhibition of either basal or (-)isoprenaline-stimulated adenylate cyclase activity was achieved with ~ 30 μ M LY-141865; maximal inhibition was achieved at concentrations ≥ 100 μ M (Fig. 2 left, and inset B). Like other dopaminergic agonists, LY-141865 decreased the maximal response to (-)isoprenaline but did not affect the molar potency of (-)isoprenaline. Fluphenazine or (-)sulpiride (dopaminergic antagonists) prevented the inhibitory effects of LY-141865 (Fig. 2 left and insets A and B). Stimulation of the β -adrenoreceptor enhances the release of α -MSH; like other dopaminergic agonists, LY-141865 diminished the ability of β -adrenergic agonists to enhance α -MSH release (Fig. 1b).

To examine the actions of LY-141865 on rat IL adenylate cyclase in the absence of β -adrenergic agonists, the drug was tested on rat IL tissue previously treated with cholera toxin. Adenylate cyclase activity was substantially greater in homogenates of cholera toxin-treated tissue than in homogenates of either fresh or control tissue (Table 1). The dopamine receptor remained functional after treatment of IL tissue with cholera toxin. Both dopamine and apomorphine (a dopaminergic agonist) markedly diminished enzyme activity in homogenates of cholera toxin-treated tissue (Table 2); (-)sulpiride or fluphenazine (dopaminergic antagonists) prevented the inhibitory effect of both dopaminergic agonists. Like dopamine and apomorphine, LY-141865 decreased adenylate cyclase activity in homogenates of cholera toxin-treated tissue (Fig. 2, right); half-maximal inhibition was achieved with 30 μ M LY-141865 (Fig. 2, inset C). The inhibitory effect of LY-141865 was attenuated in a dose-dependent manner by spiroperidol or (-)sulpiride; (-)sulpiride was 100-fold more potent than the (+)sulpiride (Fig. 2, right).

To determine whether LY-141865 is a specific D-2 agonist, the compound was tested on the dopamine-sensitive adenylate cyclase of carp retina, a tissue displaying an unusually large (four- to fivefold) dopamine-induced stimulation of adenylate cyclase activity^{9,10}. In this test system, LY-141865 displayed no activity, either as a dopaminergic agonist or as a dopaminergic antagonist (Table 3); lergotril failed to stimulate adenylate

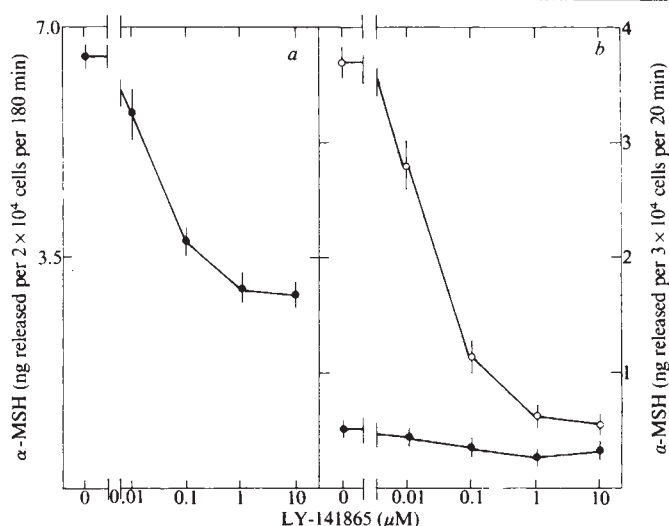


Fig. 1 LY-141865 inhibits release of α -MSH from dispersed neurointermediate lobe cells. *a*, Incubations were performed for 180 min in the presence of the indicated concentrations of LY-141865 (*trans*-($\pm$)-4,4a,5,6,7,8,8a,9-octahydro-5-propyl-2H-pyrazolo-(3,4-g)quinoline). *b*, Incubations were performed for 20 min in the absence (●) and the presence (○) of 0.1 μ M (-)isoprenaline and the indicated concentrations of LY-141865. Previously described procedures¹³ were used for the preparation of dispersed neurointermediate lobe cells, the experimental incubations and the assay of α -MSH.

cyclase activity but diminished the stimulatory effect of dopamine.

These results indicate that LY-141865 discriminates between the D-1 and D-2 receptors. Clearly LY-141865 does not interact with the D-1 receptor: it did not stimulate the D-1 receptor in the carp retina, and unlike lergotril, LY-141865 did not block this receptor²³. It is also clear that LY-141865 stimulates

Table 1 Cholera toxin increases adenylate cyclase activity in rat IL

Tissue	Adenylate cyclase activity (pmol cyclic AMP per mg protein per min)
Fresh	43 ± 2
2 h control	47 ± 21
2 h + 30 nM cholera toxin	361 ± 29

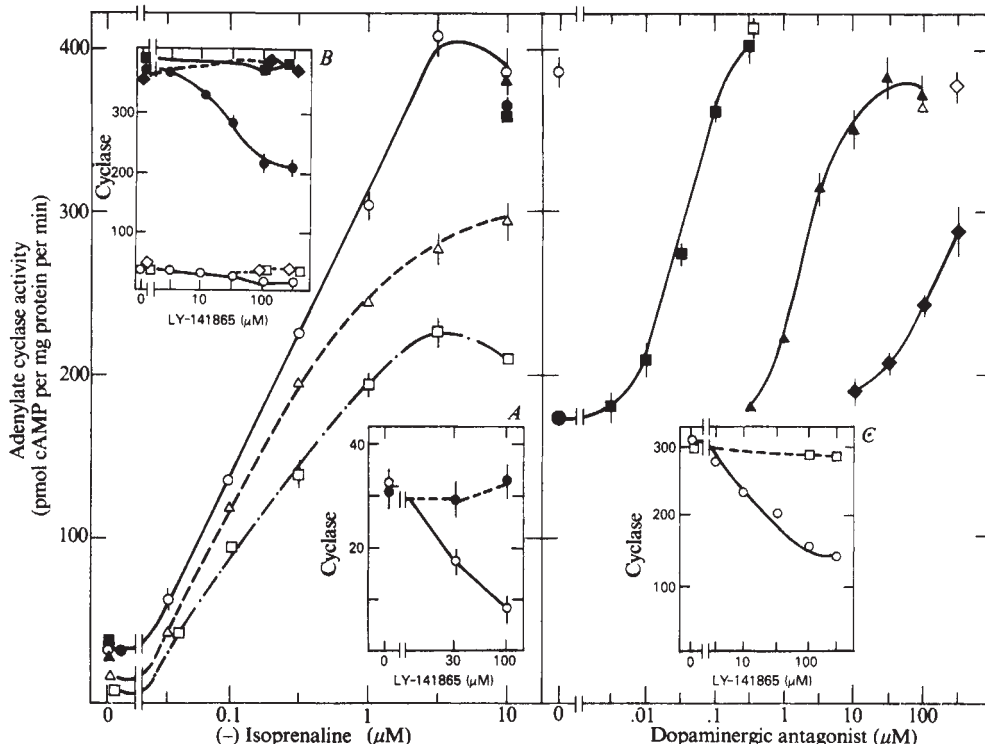
IL tissue was dissected and adenylate cyclase activity assayed as described in Fig. 2 legend. Tissue was homogenized and adenylate cyclase activity assayed either immediately after dissection (fresh) or after incubation for 2 h at 37 °C in 5 ml of RPMI 1640 medium (equilibrated with a mixture of 95% O₂ and 5% CO₂) containing 25 mM HEPES, 100 U ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin, and cholera toxin as indicated. Data represent mean $\pm$ s.e.m. for five replicate homogenates.

Table 2 Dopaminergic inhibition of adenylate cyclase activity in cholera toxin-treated rat IL

Agonist	Adenylate cyclase activity (pmol cyclic AMP per mg protein per min)		
	Antagonist		
	None	Fluphenazine (10 μ M)	(-)Sulpiride (100 μ M)
None	433 ± 14	465 ± 10	420 ± 4
Dopamine (100 μ M)	106 ± 3	453 ± 7	475 ± 9
Apomorphine (10 μ M)	150 ± 3	490 ± 7	500 ± 11

IL tissue was dissected and incubated for 2 h in the presence of 30 nM cholera toxin as described in Table 1 legend. At the end of the incubation, the tissue was homogenized and adenylate cyclase activity assayed in the presence of the indicated concentrations of test substances, as described in Fig. 2 legend. Data represent mean $\pm$ s.e.m. for four replicate samples in a single experiment.

Fig. 2 Left: LY-141865 decreases basal and (-)isoprenaline-stimulated adenylate cyclase activity in rat IL. Adenylate cyclase activity was assayed in the presence of the indicated concentrations of (-)isoprenaline alone (○) or in combination with 30 μ M LY-141865 (Δ), 100 μ M LY-141865 (□), 10 μ M fluphenazine (●), 10 μ M fluphenazine and 30 μ M LY-141865 (▲) or 10 μ M fluphenazine and 100 μ M LY-141865 (■). Inset A: Data for basal adenylate cyclase activity are replotted with a larger scale for the ordinate. Adenylate cyclase activity was assayed in the presence of the indicated concentrations of LY-141865 alone (○) or in combination with 10 μ M fluphenazine (●). Inset B: In a separate experiment, adenylate cyclase activity was assayed in the presence of the indicated concentrations of LY-141865 alone (○) or in combination with 10 μ M fluphenazine (□), 100 μ M (-)sulpiride (◇), 3 μ M (-)isoprenaline (●), 3 μ M (-)isoprenaline and 10 μ M fluphenazine (■) or 3 μ M (-)isoprenaline and 100 μ M (-)sulpiride (◆). Right: Dopaminergic antagonists reverse the inhibition by LY-141865 of adenylate cyclase activity in cholera toxin-treated IL tissue. IL tissue was treated with cholera toxin (30 nM, 2 h) as described in Table 1 and adenylate cyclase activity was determined in the absence of drugs (○) or in the presence of the indicated concentrations of spiroperidol (□), (-)sulpiride (Δ), (+)sulpiride (◇), 100 μ M LY-141865 (●) or 100 μ M LY-141865 in combination with the indicated concentrations of spiroperidol (■), (-)sulpiride (▲) and (+)sulpiride (◆). Inset C: In a separate experiment, IL tissue was treated with cholera toxin (30 nM, 2 h) and adenylate cyclase activity determined in the presence of the indicated concentrations of LY-141865 alone (○) or in combination with 100 μ M (-)sulpiride (□). Adenylate cyclase activity was determined using a previously described assay system¹⁵ containing: 80 mM, Tris-HCl (pH 7.4), 1 mM MgSO_4 , 0.8 mM EGTA, 10 mM theophylline; 0.25 mM ATP; 0.01 mM GTP; intermediate lobe tissue, 2.5–3.5 μ g of protein; and drugs as indicated in Tables 1 and 2 and Figs 1 and 2. The assay system was incubated for 10 min at 30 °C, and the reaction stopped by boiling for 2 min. Cyclic AMP concentration was determined by the method of Brown *et al.*²⁷, protein content by the method of Lowry *et al.*²⁸.



the D-2 dopamine receptor in the rat IL because (1) LY-141865 mimics the ability of dopamine and dopaminergic agonists to decrease basal or (-)isoprenaline-stimulated release of α -MSH; (2) LY-141865 mimics the inhibitory actions of dopamine and dopaminergic agonists on adenylate cyclase activity in homogenates of fresh or cholera toxin-treated tissue; and (3) the inhibitory effects of LY-141865 on adenylate cyclase activity are reversed by dopamine antagonists, including the selective D-2 antagonist, (-)sulpiride²⁴. LY-141865 is substantially more potent on intact cells of the rat IL than in the cell-free adenylate cyclase assay system (compare potency in Figs 1 and 2); the reasons for this discrepancy are unknown. The previous demonstration²¹ that LY-141865 lowers the level of serum prolactin in reserpine-treated male rats is in accord with the conclusion that LY-141865 is an agonist on the D-2 receptor. The availability of drugs selectively stimulating either the D-1 receptor (such as SKF 38393A)²⁵ or the D-2 receptor (such as LY-141865) may facilitate pharmacological analysis of the physiological responses to dopamine. Furthermore, because

LY-141865 penetrates the blood-brain barrier after its systemic administration²¹, it may be of use in the treatment of parkinsonism (as are the dopaminergic ergots²⁶).

Note added in proof: Euvard *et al.* (*Neuropharmacology* **19**, 375–386; 1980) describe two other compounds which seem to act as selective agonists on the D-2 receptor.

Received 9 April; accepted 19 May 1981.

- Cools, A. R. & Van Rossum, J. M. *Psychopharmacologia* **45**, 243–254 (1976).
- Titeler, M., Weinreich, P., Sinclair, D. & Seeman, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1153–1156 (1978).
- Costal, B. & Naylor, R. J. *Life Sci.* **28**, 215–229 (1981).
- Spano, P. F. *et al.* in *Dopaminergic Ergot Derivatives and Motor Function* (eds Fuxe, K. & Calne, D. B.) 159–171 (Pergamon, Oxford, 1979).
- Kebabian, J. W. & Calne, D. B. *Nature* **277**, 93–96 (1979).
- Brown, E. M., Carroll, R. J. & Aurbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1113–1117 (1977).
- Attie, M. *et al.* *Endocrinology* **107**, 1776–1781 (1980).
- Brown, E. M. *et al.* *Molec. Pharmacol.* **18**, 335–340 (1980).
- Redburn, D., Clement-Cormier, Y. C. & Lam, D. M. K. *Life Sci.* **27**, 23–31 (1980).
- Watling, K. J., Dowling, J. E. & Iversen, L. L. *Nature* **281**, 578–580 (1979).
- MacLeod, R. M. in *Frontiers in Neuroendocrinology* Vol. 4 (eds Martini, L. & Ganong, W. F.) 169–194 (Raven, New York, 1976).
- Caron, M. G. *et al.* *J. biol. Chem.* **253**, 2244–2253 (1978).
- Munemura, M., Eskay, R. L. & Kebabian, J. W. *Endocrinology* **106**, 1795–1803 (1980).
- Munemura, M. *et al.* *Endocrinology* **107**, 1676–1683 (1980).
- Cote, T. E. *et al.* *Endocrinology* **107**, 108–116 (1980).
- Kebabian, J. W., Calne, D. B. & Kebabian, P. R. *Commun. Psychopharmacol.* **1**, 311–318 (1977).
- Schmidt, M. J. & Hill, L. E. *Life Sci.* **20**, 789–798 (1977).
- Pieri, L., Keller, H. H., Burkard, W. & DaPrada, M. *Nature* **272**, 278–280 (1978).
- Mowles, T. F., Burghardt, B., Burghardt, C., Charnecki, A. & Sheppard, H. *Life Sci.* **22**, 2103–2112 (1978).
- Miller, R. J. & Iversen, L. L. *Naunyn-Schmiedeberg's Archs Pharmacol.* **282**, 213–216 (1974).
- Bach, N. J. *et al.* *J. med. Chem.* **23**, 481–491 (1980).
- Cote, T. E., Grewe, C. W. & Kebabian, J. W. *Endocrinology* **108**, 420–426 (1981).
- Watling, K. J. & Dowling, J. E. *J. Neurochem.* **36**, 559–568 (1981).
- Spano, P. F., Stefanini, E., Trabucchi, M. & Fresia, P. in *Sulpiride and Other Benzamides* (eds Spano, P. F. *et al.*) 11–31 (Raven, New York, 1979).
- Setler, P. E., Sarau, H. M., Zerkle, C. L. & Saunders, H. L. *Eur. J. Pharmacol.* **50**, 419–430 (1978).
- Thorner, M. O., Fluckiger, E. W. & Calne, C. B. *Bromocriptine: A Clinical and Pharmacological Review*, 1–189 (Raven, New York, 1980).
- Brown, B. L. *et al.* *Biochem. J.* **121**, 561–562 (1971).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).

Table 3 LY-141865 fails to stimulate adenylate cyclase of fish retina

Dopamine concentration (μ M)	Adenylate cyclase activity (pmol cyclic AMP per mg protein per min)		
	Test compound		
	None	LY-141865 (100 μ M)	Lergotriole (100 μ M)
0	11.1 $\pm$ 0.7	9.4 $\pm$ 0.8	6.2 $\pm$ 1.1
10	35.5 $\pm$ 1.7	36.3 $\pm$ 2.4	12.5 $\pm$ 0.7

Retinal tissue was obtained from goldfish purchased at a local pet store. Adenylate cyclase activity was assayed as described in Fig. 2 legend. Data represent mean $\pm$ s.e.m. for four replicate samples obtained in a single experiment.

Trypanocidal activity of daunorubicin and related compounds

J. Williamson & T. J. Scott-Finnigan

Division of Parasitology, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

M. A. Hardman & J. R. Brown*

Department of Pharmaceutical Chemistry, School of Pharmacy, Leicester Polytechnic, PO Box 143, Leicester LE1 9BH, UK

The anthracycline antibiotic daunorubicin is, *in vitro*, one of the most potent trypanocidal agents reported¹; it permanently abolishes the infectivity of African trypanosomes at less than nanomolar concentrations. In contrast, other anthracyclines such as doxorubicin and nogalamycin must be present in millimolar concentrations to achieve the same effect (Table 1). Despite its uniquely high activity *in vitro*, daunorubicin is totally inactive against trypanosomes in infected rodents¹. An investigation of this lack of *in vivo* activity suggested that uptake into trypanosomes is only transient². We therefore tried to prolong exposure of trypanosomes to the drug by coupling it to a carrier macromolecule. The use of such a carrier should delay removal of the drug from the blood and simultaneously exploit the endocytotic properties of trypanosomes. Ferritin and albumin were chosen as potential carriers as these proteins are endocytosed by trypanosomes^{3,4}. We report here that when the drug was attached by a stable linkage, activity was not maintained, but when attached by a labile covalent linkage the preparation was active both *in vitro* and *in vivo*, indicating the desirability of testing other macromolecules, such as dextran⁵, as carriers for trypanocidal drugs.

As none of the daunorubicin analogues tested (see Table 1, and Fig. 1) reproduce the high *in vitro* trypanocidal activity of daunorubicin, the loss of activity of this drug *in vivo* must be overcome. We first tested a DNA-daunorubicin complex⁶ but it was inactive. As this product acts as a slow-release form of drug⁷ rather than a true carrier, covalent attachment of drug to possible carriers was then investigated. Bovine serum albumin (BSA) and ferritin were chosen as carriers because of the likelihood of their endocytosis by trypanosomes^{3,4}. The first albumin conjugate was prepared by oxidative cleavage of the amino sugar of the drug to yield an aldehyde which was reacted with the protein to give a Schiff's base—this was then reduced with borohydride⁸. The product was almost inactive (Table 1), possibly due to destruction of the amino sugar unit or to the stability of the linkage. Therefore a second conjugate⁸ was prepared using glutaraldehyde to form Schiff's bases simultaneously with the amino group of daunorubicin and a free amino group of the protein. This product retained activity *in vitro* (Table 1) and was significantly active *in vivo* (Table 2). As Schiff's bases are susceptible to hydrolysis, the rate of drug release from the conjugate in aqueous solution (pH 7.4) at 37 °C was monitored (by size-exclusion HPLC); release seemed to be a first-order process with a half life of ~100 h.

The retention of activity only when the carrier-binding linkage is labile suggests that the drug must be released before it can exert its effects. Because of the success of albumin as a carrier, we also prepared ferritin conjugates in which daunorubicin was linked by the glutaraldehyde method⁸. These were also found to have high activity *in vitro* and *in vivo* (Tables 1 and 2). These glutaraldehyde-coupled daunorubicin-albumin and daunorubicin-ferritin complexes were the first daunorubicin preparations to clear trypanosomes from the blood of infected mice for an extended period. For 20–30 h after treatment, the proportion of abnormal trypanosomes rose to almost 100%, the highest proportions coinciding with a sharp fall in parasitaemia.

* Present address: Department of Pharmaceutical Chemistry, Faculty of Pharmaceutical Sciences, Sunderland Polytechnic, Chester Road, Sunderland SR1 3SD, UK.

Table 1 *In vitro* activity of daunorubicin and related compounds against *Trypanosoma rhodesiense*

	Loss of motility	Loss of infectivity Temporary	Permanent
a Daunorubicin and its analogues			
(1) Daunorubicin hydrochloride	4	11	10
(2) Doxorubicin hydrochloride	8	>10	4
(3) 4-Demethoxydaunorubicin hydrochloride	5	5	<3
(4) 4'-Deoxydaunorubicin hydrochloride	5	6	3
(5) Marcellomycin	10	10	4
(6) <i>N</i> -Trimethyl-daunorubicin chloride	<3	3	<3
(7) Doxorubicin-14-stearate hydrochloride	3	<3	3
(8) AD32	3	5	<3
(9) Dibenzyl-daunorubicin	<3	<3	<3
(10) Dihydrochloride	4	7	3
(11) Hydrochloride	8	>9	3
(12) Dihydrochloride	7	7	3
(13) Dihydrochloride	7	10	4
b Daunorubicin-protein complexes			
Daunorubicin-BSA (P) (22 µg drug per mg)	4	4	<4
Daunorubicin-BSA (G) (18 µg drug per mg)	<4	>10	>10
Daunorubicin-ferritin (G) (55 µg drug per mg)	4	10	6
Daunorubicin-ferritin (C) (38 µg drug per mg)	<3	>9	4

In these experiments we used the monomorphic Liverpool strain of *T. rhodesiense*¹¹. For loss of motility and infectivity, the values given are maximum titres ($\log_{10} M^{-1}$) producing the effect; drugs with titres <3 are considered inactive. The assay method is given elsewhere¹. P denotes that the periodate oxidation coupling method⁸ was used, G that the glutaraldehyde coupling method⁸ was used, and C that the daunorubicin is sequestered in the iron core. Compound sources: 1–4, Dr F. Arcamone; 5, Dr W. T. Bradner; 6, Dr E. M. Acton; 7, was prepared by an adaptation of the method described for doxorubicin-14-octanoate¹²; 8, Dr M. Israel; 9, The Drug Synthesis and Chemistry Branch, Division of Cancer Treatment, NCI; 10, Dr K. C. Murdock.

Table 2 Activity of daunorubicin-carrier complexes in mice infected with *T. rhodesiense*

Compound	Dose of drug or drug component (mg per kg)	Proportion of mice cleared of infection
Daunorubicin	30	0/4*
Daunorubicin-DNA	30	0/5*
Daunorubicin-BSA (G)	30	7/11*
	15	11/13*
	7.5	10/11
	3.8	0/11
Daunorubicin-ferritin (G)	1.9	0/11
	30	8/10
Daunorubicin-ferritin (G)	15	11/13
Daunorubicin-ferritin (C)	15	0/13

Mice with a patent parasitaemia were given a single intraperitoneal dose of the compound being tested 24–48 h after infection. Control mice were uniformly dead 4–5 days after infection. Generally, the amounts of the conjugates available were small and consequently the numbers of mice used in the tests were limited. G denotes conjugation by the glutaraldehyde method⁸; C denotes non-covalent binding of drug in the ferritin iron core.

* Some deaths due to drug toxicity at this dose.

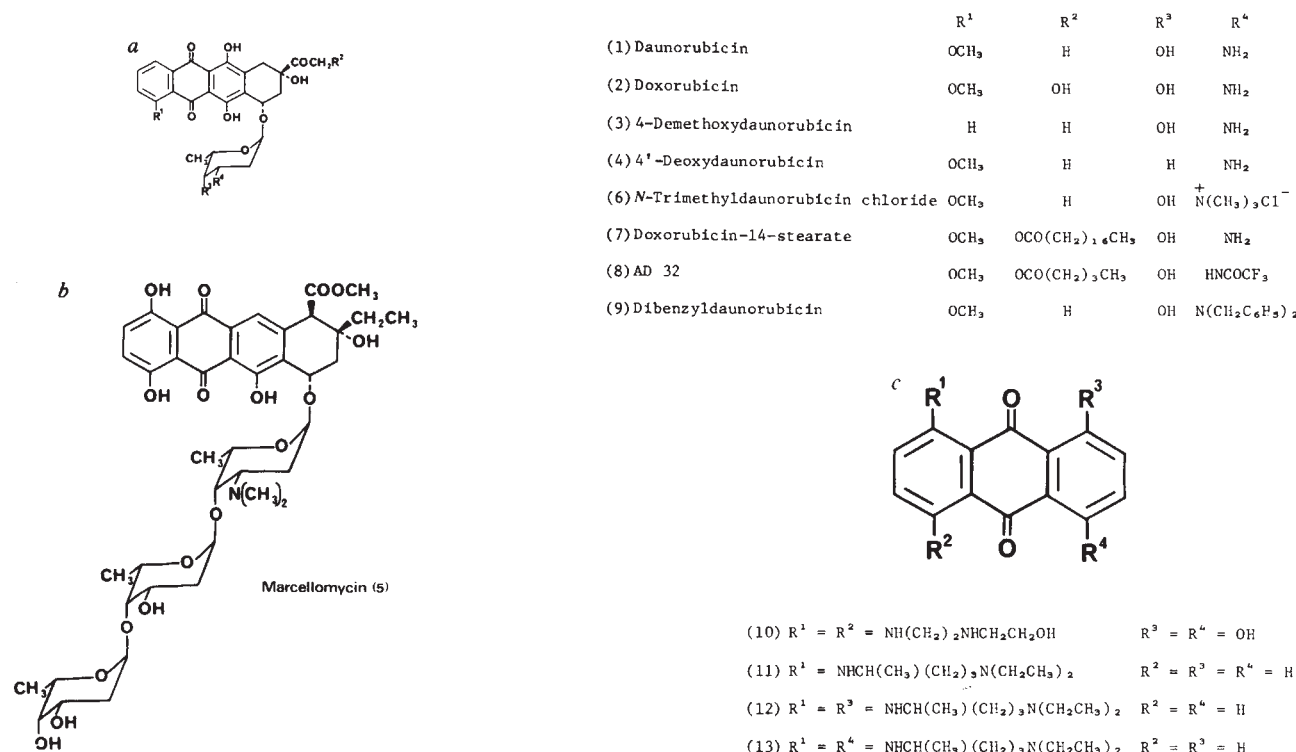


Fig. 1 Structures of the compounds tested for trypanocidal activity.

Fluorescence microscopy showed that the parasites took up the complexed drug progressively for up to 12 h after treatment and retained it thereafter, unlike the transient uptake peaking at 1.5–2 h of the free drug. This altered pattern of uptake suggests that drug is taken up as the drug-protein complex; the preparation thus seems to act as a true endocytotic carrier.

Ferritin contains an iron (Fe³⁺) core surrounded by 24 subunits of apoferritin and can be reconstituted from apoferritin in the presence of Fe³⁺ or Fe²⁺ and oxidizing agents⁹. As daunorubicin forms a chelate with iron, preparation of a daunorubicin-ferritin complex by sequestration of the drug in the iron core of ferritin was attempted by reconstitution of ferritin from apoferritin and Fe²⁺ with a weak oxidizing agent, in the presence of daunorubicin. A high uptake of drug was achieved and despite the non-covalent nature of the complex, the drug was not released on gel chromatography of the complex in 1M NaCl or in 2% SDS in 8M urea; the product was however less active than the covalently bound preparation and showed no *in vivo* activity (Tables 1 and 2).

The need for new drugs active against African trypanosomiasis, both human and animal, is widely recognized¹⁰. Our results clearly demonstrate that daunorubicin has considerable potential as a trypanocide as: (1) it shows unparalleled trypanocidal activity *in vitro*; (2) it can be conjugated to macromolecular carriers to give soluble complexes which retain activity *in vivo*; (3) it is structurally distinct from existing trypanocides and so should lack cross-resistance with these drugs; and (4) it has considerable capacity for structural variation.

We thank Dr Diane J. McLaren for providing electron micrographs, Miss Jane C. Marsh for technical assistance and the Wellcome Trust for support (to M.A.H.). We also thank the suppliers of the daunorubicin analogues listed in Table 1.

- Hulhoven, R. & Harvengt, C. *Biomedicine* **33**, 44 (1980).
- Hurwitz, E., Levy, R., Wilcheck, M., Arnon, R. & Sela, M. *Cancer Res.* **35**, 1175 (1975).
- Crichton, R. *Angew. Chem.* **12**, 57 (1973).
- The African Trypanosomiasis* Tech. Rep. Ser. 635 (WHO/FAO, Geneva, 1979).
- Yorke, W., Adams, A. R. D. & Murgatroyd, F. *Ann. trop. Med. Parasit.* **23**, 501 (1929).
- Arcamone, F. *et al. J. med. Chem.* **17**, 335 (1974).

Interactions of daunomycin and melanotropin-daunomycin with DNA

Gary Wieseahn*, Janos M. Varga†
& John E. Hearst*

* Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, California 94720, USA

† Departments of Biochemistry and Dermatology, Yale Medical School, Yale University, New Haven, Connecticut 06520, USA

The specific recognition of fluorescein-melanotropin and melanotropin-ferritin-fluorescein conjugates by melanotropin (MSH) receptors on the surface of murine melanoma cells suggested the possibility of specific site-directed chemotherapy of melanomas¹⁻³. An MSH-daunomycin conjugate has been shown to be more toxic to mouse melanoma cells than free daunomycin, and, at the same concentration, it was not toxic to mouse 3T3 fibroblast cells. Unconjugated daunomycin was equally toxic to both cell lines⁴. We report here investigation of the interactions of purified DNA with free daunomycin and with MSH-daunomycin conjugate using spectrophotometry and also an agarose gel assay that detects the unwinding of DNA due to ligand binding. We found an association constant of $4.8 \times 10^6 \text{ M}^{-1}$ and an unwinding angle of $10 \pm 1.5^\circ$ for daunomycin, but no detectable binding and an association constant of $< 4.8 \times 10^3 \text{ M}^{-1}$ for MSH-daunomycin conjugate. We conclude that the cytotoxicity of MSH-daunomycin is not due to its binding to DNA.

Received 7 November 1980; accepted 9 June 1981.

- Williamson, J. & Scott-Finnigan, T. J. *Antimicrob. Agents Chemother.* **13**, 735 (1978).
- Brown, J. E., Brown, J. R., & Williamson, J. *Pharm. Pharmacol.* (submitted).
- Brown, K. N., Armstrong, J. A. & Valentine, R. C. *Exptl Cell Res.* **39**, 129 (1965).
- Fairlamb, A. H. & Bowman, I. B. R. *Exptl Parasit.* **43**, 353 (1977).
- Bernstein, A. *et al. J. natn. Cancer Inst.* **60**, 379 (1978).
- Trouet, A., Deprez-de Campanere, D., de Smedt-Malengreux, M. & Atassi, G. *Eur. J. Cancer* **10**, 405 (1974).

The cytotoxicity of daunomycin is thought to be due to its strong interaction with cellular DNA^{5,6}, thereby interfering with both RNA and DNA synthesis^{7,8}. The experiments described here were designed to determine whether the interaction of daunomycin with DNA is affected by the covalent coupling of daunomycin to MSH. The change in the absorption spectrum of daunomycin as increasing amounts of supercoiled Col EI DNA are added is shown in Fig. 1A. With increasing concentration of DNA, the absorption maximum shifts gradually from 480 nm to 505 nm and the absorption of daunomycin decreases towards a limit that represents the spectrum of daunomycin in the fully complexed form. There appears to be an isosbestic point at ~545 nm. Assuming independent binding sites, Scatchard analysis⁹ will result in the following expression:

$$\frac{r}{P} = K_{APP} \times (B_{APP} - r)$$

Following the nomenclature of Bresloff and Crothers¹⁰, r is the ratio of the concentration of bound ligand to the total concentration of DNA base pairs, K_{APP} is the apparent association constant for the interaction of DNA and ligand, B_{APP} is the

number of apparent binding sites for ligand per base pair and P the concentration of free ligand.

A Scatchard plot calculated from the spectrophotometric results on daunomycin binding is shown in Fig. 1B. The plot reveals a K_{APP} of $4.8 \times 10^6 \text{ M}^{-1}$ and a B_{APP} of 0.24 for the binding of daunomycin to DNA. These values are in reasonable agreement with published values of $6.6 \times 10^6 \text{ M}^{-1}$ and 0.36 for daunomycin binding to calf thymus DNA in 0.1 M Tris-HCl (pH 7.0)⁶.

In contrast, Fig. 1C shows that minimal spectral change (only a very slight initial depression in the absorption profile) occurred when Col EI DNA was added to the MSH-daunomycin conjugate. Thus no binding was detected by this method.

An additional test for DNA binding is the unwinding assay in which supercoiled DNA is relaxed with nicking-closing enzyme in the presence of compounds that bind DNA. The result is a complex between drug and covalently closed DNA with the DNA in a relaxed conformation. On removal of the drug the DNA becomes superhelical, and the number of superhelical turns can be determined by agarose gel electrophoresis¹¹.

Figure 2A illustrates a comparison of DNA-unwinding assays with daunomycin (lanes a-f) and MSH-daunomycin conjugate (lanes g-l) at comparable drug/DNA ratios. Daunomycin clearly has a large effect on the superhelical density of DNA,

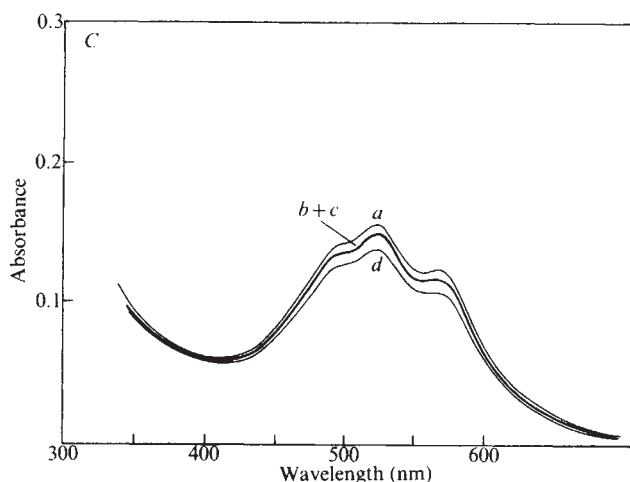
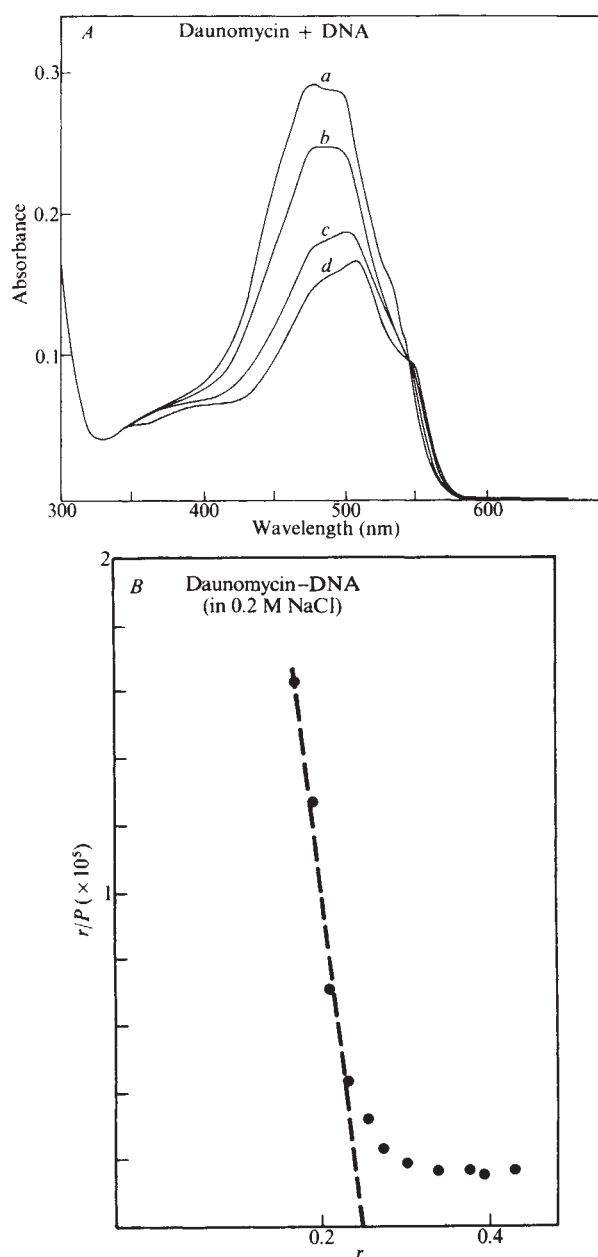


Fig. 1 A, effect of supercoiled colicin EI (Col EI) plasmid DNA, isolated from *E. coli* JC 411 *thy* (Col EI), on the absorption spectrum of daunomycin. The incubation mixture contained 2.89×10^{-5} M daunomycin in 0.04 M Tris-HCl, pH 8.0, 5 mM NaOAc, 0.5 mM EDTA, 0.2 M NaCl. Absorbance was measured using a 1-cm path length quartz cuvette in a Cary 118 spectrophotometer, at the following DNA concentrations (M): curve a, zero; b, 2.57×10^{-5} ; c, 1.03×10^{-4} ; d, 2.14×10^{-4} . These concentrations are in moles of DNA base pairs per litre. B, Scatchard plot of the data in A, plotted according to the method of Muller and Crothers¹⁶. We determined the ratio of free to bound daunomycin at intermediate points by assuming the absorption spectrum is a linear combination of the free and fully complexed daunomycin spectra. The absorbance at 470 nm was used for this calculation as at this wavelength the difference between free and bound daunomycin is greatest. For example, we assumed that the ratio of free to bound daunomycin is 2:1 when the absorption at 470 nm is one-third of the way from the absorption of free daunomycin to the absorption of fully complexed daunomycin. Knowing the total concentration of daunomycin, it is then simple to calculate the concentrations of free and bound daunomycin. The absorption data were corrected for the dilution due to addition of DNA solution. C, effect of supercoiled Col EI DNA on the absorption spectrum of MSH-daunomycin conjugate. The incubation mixture contained 2.11×10^{-5} M conjugate in the buffer described in A. The absorbance was measured at the following DNA concentrations (M): curve a, 0; b, 1.03×10^{-5} ; c, 2.57×10^{-5} ; d, 3.48×10^{-4} . The slight depression in the absorbance spectrum is due to dilution of the conjugate on addition of the DNA. The conjugate was prepared as described by Varga *et al.*⁴.

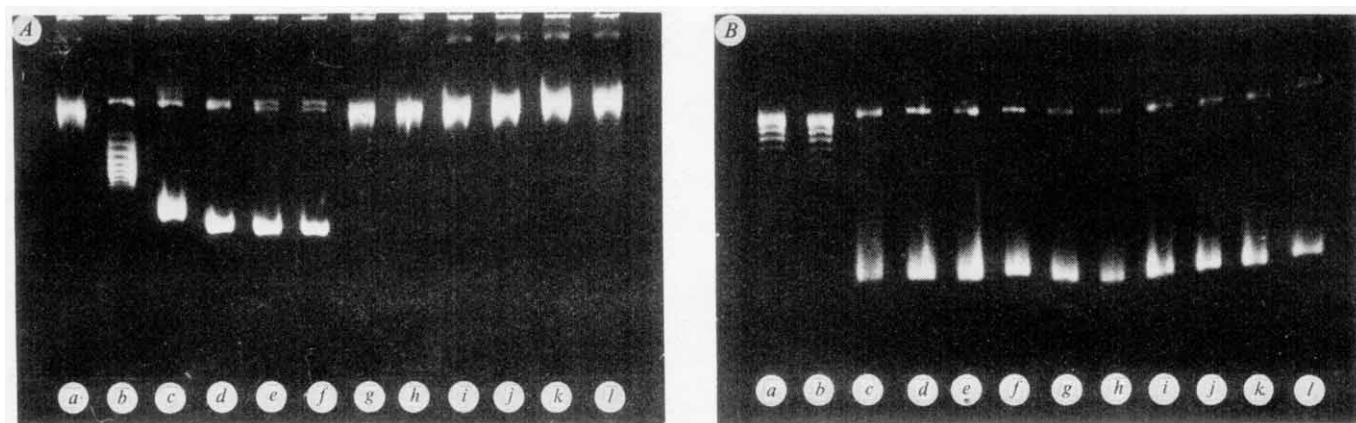


Fig. 2 *A*, agarose slab gel of Col EI DNA (4.3×10^{-5} M) relaxed with nicking-closing enzyme in the presence of increasing amounts of daunomycin (lanes *a-f*) or MSH-daunomycin conjugate (lanes *g-l*). Lanes: *a*, no daunomycin present; *b*, 2.05×10^{-6} M; *c*, 4.10×10^{-6} M; *d*, 6.14×10^{-6} M; *e*, 8.19×10^{-6} M; *f*, 1.02×10^{-5} M; *g*, no conjugate present; *h*, 1.49×10^{-6} M; *i*, 3.00×10^{-6} M; *j*, 4.49×10^{-6} M; *k*, 5.98×10^{-6} M; *l*, 7.48×10^{-6} M. The unwinding assay was performed as described elsewhere¹¹. *B*, agarose slab gel of Col EI DNA (1.3×10^{-5} M) reacted with nicking-closing enzyme in the presence of increasing amounts of daunomycin-MSH conjugate. Lanes: *a*, no conjugate present; *b*, 7.50×10^{-6} M; *c*, 2.25×10^{-5} M; *d*, 3.75×10^{-5} M; *e*, 5.25×10^{-5} M; *f*, 6.75×10^{-5} M; *g*, 8.25×10^{-5} M; *h*, 9.75×10^{-5} M; *i*, 1.12×10^{-4} M; *j*, 1.28×10^{-4} M; *k*, 1.42×10^{-4} M; *l*, 1.58×10^{-4} M. If we assume that the binding of one molecule of conjugate would unwind DNA by 10° as does the binding of unconjugated daunomycin, then the binding of 1% of the conjugate molecules (at a base pair/conjugate ratio of 1:1) would cause about a two-band shift in the DNA distribution on the agarose gel ($64 \times 10^\circ = 640^\circ$ or 1.8 superhelical turns). This two-band shift should be easily detected on the gel. As no shift was observed, we conclude that the ratio of bound to unbound conjugate is <0.01 . This ratio can be compared with the ratio of 12 bound daunomycin molecules per unbound daunomycin molecule calculated from the Scatchard equation at low values of r . Thus, the bound/unbound ratio for the conjugate is at least three orders of magnitude less than for daunomycin, and $K_{APP} < 4.8 \times 10^3$ M $^{-1}$ for the conjugate.

whereas the conjugate has no detectable effect. Using the binding constant 4.8×10^6 M $^{-1}$ and assays such as are shown in Fig. 2*A*, we have determined an unwinding angle of $10 \pm 1.5^\circ$ for each bound daunomycin molecule following the method previously described¹². This again agrees with the value of $12 \pm 3.3^\circ$ (corrected to an ethidium unwinding angle of 28°) published previously by Waring¹².

As shown in Fig. 2*B*, increasing the concentration of MSH-daunomycin conjugate in an attempt to detect low-level binding to DNA does not result in the production of a population of DNA molecules of intermediate superhelical density as in lanes *b* and *c* of Fig. 2*A*. It results instead in inhibition of nicking-closing activity (lanes *c-l*, Fig. 2*B*). The highest concentration of MSH-daunomycin conjugate that produced fully relaxed Col EI DNA corresponded to one conjugate molecule per base pair ($\sim 6,400$ conjugate molecules per Col EI DNA molecule). The inhibition of nicking-closing enzyme by MSH-daunomycin conjugate seems to be due to an interaction between the nicking-closing enzyme and the conjugate because increased concentrations of enzyme require increased concentrations of conjugate for inhibition. This inhibition seems to be mainly due to MSH and occurs at rather high concentrations of conjugate (50 μ M) (data to be published elsewhere).

We could not detect any binding of MSH-daunomycin conjugate to DNA which results in a shift of the absorption spectrum of daunomycin nor any which distorts the DNA helix enough to be detected in a DNA-unwinding assay. We conclude from the DNA-unwinding assay that the association constant for binding of the conjugate to DNA is $<4.8 \times 10^3$ M $^{-1}$. This lack of binding is consistent with the observation of Zunino *et al.*⁶, who showed that the *N*-acetylation of the sugar ring of daunomycin reduces the affinity for DNA by more than two orders of magnitude. The addition of the large MSH molecule to the split sugar ring (which also results in the removal of the primary amino group) might be expected to decrease the affinity of daunomycin to DNA even more.

The killing of mouse melanoma cells by melanotropin-daunomycin conjugate must therefore be explained by a mechanism other than direct binding of the conjugate to DNA. It has recently been shown that adriamycin, a structural analogue of daunomycin, perturbs the cell membrane by lowering the temperature of gel-to-liquid transition¹³. Thus the site of

action of melanotropin-daunomycin might be the cell membrane. However, the effects of transglutaminase inhibitors on the toxicity of the conjugate strongly suggest an effect on internalization. It has been demonstrated that inhibitors of transglutaminase inhibit receptor-mediated endocytosis of polypeptide hormones¹⁴. If endocytosis of melanotropin-daunomycin is essential for toxicity, we could expect that inhibitors of transglutaminase would prevent the toxic effects of the conjugate. We have, indeed, found that ammonium acetate (1 mM), methylamine (0.1 mM) and chloroquine (10^{-7} M) all reduce the toxic effects of MSH-daunomycin by about 50%¹⁵. As direct binding of the conjugate to DNA does not explain its toxic effect, we must therefore consider two alternative explanations—that the free and conjugated toxins kill the cells by different mechanisms, or that the internalized MSH-daunomycin is metabolized to a form that can bind to DNA and thereby interfere with transcription and replication. We think the second mechanism more likely.

This work was supported by the Assistant Secretary for Environment, Office of Environmental Research and Development, Biomedical and Environmental Research Division of the US Department of Energy under contract no. W-7405-ENG-48, by grant no. CA-26081 from the NCI and by Research Career Development Award no. 5K04 AI00154.

Received 2 February; accepted 28 May 1981.

- Varga, J. M., DiPasquale, A., Pawelek, J., McGuire, J. S. & Lerner, A. B. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1590-1593 (1974).
- Varga, J. M., Moellmann, G., Fritsch, P., Godawska, E. & Lerner, A. B. *Proc. natn. Acad. Sci. U.S.A.* **73**, 559-562 (1976).
- Lerner, A. B., Moellmann, G., Varga, J. M., Halaban, R. & Pawelek, J. *Cold Spring Harb. Conf. on Cell Proliferation* **6**, 187-197 (1979).
- Varga, J. M., Asato, N., Lande, S. & Lerner, A. B. *Nature* **267**, 56-58 (1977).
- Penco, S. *Chimica Ind., S. Paulo* **50**, 908 (1968).
- Zunino, F., Gambetta, R., DiMarco, A. & Zaccara, A. *Biochim. biophys. Acta* **277**, 489-498 (1972).
- Ward, D. C., Reich, E. & Goldberg, I. H. *Science* **149**, 1259-1263 (1965).
- Hartmann, G., Goller, H., Koschel, K., Kersten, W. & Kersten, H. *Biochem. Z.* **341**, 126-128 (1964).
- Scatchard, G. *Ann. N. Y. Acad. Sci.* **51**, 660-672 (1949).
- Bresloff, J. L. & Crothers, D. M. *J. molec. Biol.* **95**, 103-123 (1975).
- Wiesehahn, G. & Hearst, J. E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2703-2707 (1978).
- Waring, M. *J. molec. Biol.* **54**, 247-279 (1970).
- Tritton, T. R., Murphree, S. A. & Sartorelli, A. C. *Biochem. biophys. Res. Commun.* **84**, 802-808 (1978).
- Davies, P. J. A. *et al. Nature* **283**, 162-167 (1980).
- Varga, J. M., Asato, N., Wiesehahn, G. & Hearst, J. E. *Pigment Cell* **6** (in the press).
- Muller, W. & Crothers, D. M. *J. molec. Biol.* **35**, 251-290 (1969).

Enteropancreatic circulation of digestive enzymes does not exist in the rat

Gerhard Rohr*, Horst Kern† & George Scheele‡

* Medizinische Universitätsklinik and † Institut für Anatomie und Zellbiologie, Universität, D-3550 Marburg/Lahn, FRG

‡ Department of Cell Biology, Rockefeller University, New York, New York 10021, USA

The hypothesis of an enteropancreatic circulation¹⁻³, that digestive enzymes are absorbed intact across the intestinal epithelium, are circulated through the bloodstream and resecreted by the exocrine pancreas, has generated considerable discussion⁴⁻⁶ and has been accepted by a number of medical scientists⁷⁻⁹. However, the data presented in support of this hypothesis do not demonstrate with certainty that labelled proteins introduced into the bathing medium^{1,3}, intestinal tract^{1,2} or bloodstream² appear unmodified in pancreatic juice. We have now critically examined that part of the hypothesis which proposes that pancreatic secretory proteins are taken up from the bloodstream and secreted intact into the pancreatic juice. Injection of ³⁵S-methionine-labelled rat pancreatic amylase into the blood circulation of unanaesthetized rats resulted in the appearance of radioactivity in all exocrine pancreatic proteins and showed a distribution which was indistinguishable from that following the intravenous administration of ³⁵S-methionine. These findings are not consistent with the direct transport of amylase from the blood circulation to the pancreatic ductal lumen, but are consistent with metabolic degradation of the polypeptide probe and reutilization of the free ³⁵S-methionine for the synthesis and secretion of the entire complement of pancreatic exocrine proteins. In contrast, analysis of bile by similar methods indicated the presence of a haematobiliary pathway, which accounted for the excretion of 0.8% of exocrine pancreatic proteins injected into the blood circulation.

³⁵S-methionine-labelled rat pancreatic amylase was injected into the blood circulation of nonanaesthetized rats and the appearance of radioactivity in pure pancreatic juice was followed with time and further analysed by two-dimensional isoelectric focusing/SDS-gel electrophoresis followed by fluorography. Rats (250–300 g) were prepared for study as follows. The pancreatic duct was cannulated with a 0.6-mm PVC catheter, fixed with a surgical ligature and routed subcutaneously to the nape of the neck. Bile flow into the intestinal tract was re-established during the surgical procedure using a short cannula of the same size. After the operation rats were maintained in restraining cages, which allowed exercise and free access to food and water. Rats were studied 6–18 h after the operation when pancreatic secretion rates were stable, usually between 450 and 500 $\mu\text{l h}^{-1}$. As the control, ³⁵S-methionine was injected into the tail vein of nonanaesthetized rats.

Trichloroacetic acid (TCA)-insoluble radioactivity first appeared in pancreatic juice 1.5 h after the injection of ³⁵S-labelled amylase (Fig. 1). Radioactivity increased slowly thereafter and peaked at 6 h. In contrast, after the injection of ³⁵S-methionine, TCA-insoluble radioactivity appeared in pancreatic juice within 30 min and peaked at 2 h. In both conditions, we observed no difference between TCA-insoluble and total radioactivity, indicating that free methionine was not released into the ductal fluid. These findings clearly differ from those of previous investigators who claimed¹ that pancreatic proteins were transported from the interstitial space of the exocrine pancreas to the duct lumen within 15 min.

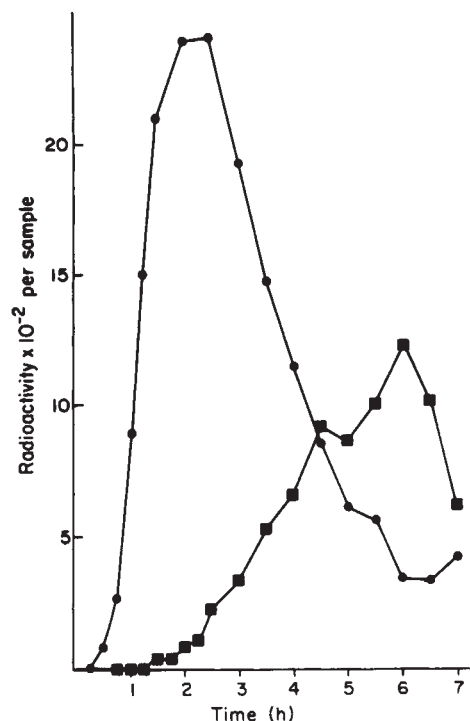
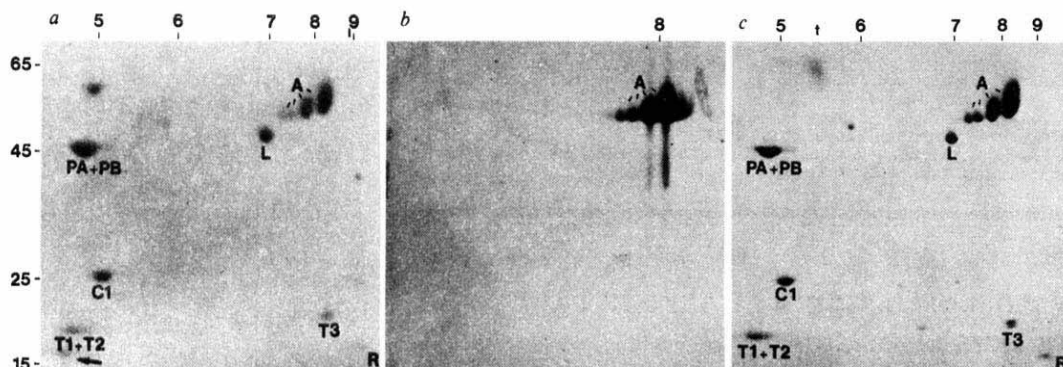


Fig. 1 Appearance of radioactivity in pancreatic juice as a function of time after intravenous administration of ³⁵S-methionine-labelled rat pancreatic amylase (2×10^6 c.p.m.) compared with ³⁵S-methionine (1.5×10^7 c.p.m.). TCA-insoluble radioactivity is shown on the ordinate for 100- μl samples collected after the injection of amylase (■) or for 10- μl samples after the injection of ³⁵S-methionine (●). Radiolabelled amylase was prepared as follows. Exocrine pancreatic proteins contained in rat pancreatic lobules were labelled with ³⁵S-methionine for 3 h as previously described for the dog pancreas¹². A crude zymogen granule fraction was isolated by differential sedimentation and proteins contained within it were liberated by treatment with 0.2 M NaHCO₃, pH 8.5. ³⁵S-methionine-labelled amylase was precipitated with shellfish glycogen by the procedure of Schramm and Loyter¹³. The glycogen amylase precipitate was washed twice as described previously¹¹ and solubilized in normal saline.

Radioactive protein(s) injected into the blood circulation and those appearing in pure pancreatic juice were separated by two-dimensional isoelectric focusing/SDS-gel electrophoresis using the procedure of Bieger and Scheele¹⁰ and analysed by Coomassie blue staining and fluorography. Figure 2a shows the Coomassie blue-staining pattern of proteins found in rat pancreatic juice. Most of these proteins have been identified according to actual or potential enzyme activity (J. Schick, H. Kern and G. Scheele, unpublished observations) by methods previously described for the guinea pig pancreas¹¹. Figure 2b shows the distribution of radioactive proteins injected into the blood circulation and confirms that the injected material contained exclusively ³⁵S-methionine-labelled rat pancreatic amylase. Figure 2c shows that radioactivity is distributed among all exocrine proteins in pancreatic juice at the 6-h time point (maximal radioactivity). After injection of ³⁵S-methionine, pancreatic juice showed a similar distribution of radioactive proteins as judged by both fluorography and liquid scintillation spectroscopy following excision of Coomassie blue-stained spots and solubilization of the acrylamide pieces (data not shown). Samples of pancreatic juice collected earlier, 3–6 h after the injection of radiolabelled amylase, also showed a similar distribution of radioactivity among all secreted proteins.

The per cent radioactivity precipitated with shellfish glycogen allowed us to quantify the distribution of amylase among

Fig. 2 Analysis of proteins injected into the blood circulation or collected in pancreatic juice by two-dimensional isoelectric focusing/SDS-gel electrophoresis. *a* Shows the Coomassie blue-staining pattern of proteins contained in rat pancreatic juice; *b* shows a fluorograph of ^{35}S -methionine-labelled rat pancreatic proteins precipitated with shellfish glycogen and subsequently injected into the blood



circulation; and *c* shows a fluorograph of proteins contained in pancreatic juice 6 h after injection of the radiolabelled material presented in *b*. After intravenous administration of ^{35}S -methionine, the distribution of radioactive proteins secreted into the pancreatic juice showed a fluorographic pattern similar to that observed in *c*. Each of the gels in panels *a*–*c* were impregnated with diphenyloxazole as required for fluorography. Proteins are labelled according to their actual or potential enzyme activities: A, amylase; L, lipase; PA, procarboxypeptidase A; PB, procarboxypeptidase B; C, chymotrypsinogen; T, trypsinogen; and R, ribonuclease. The arrow in *a* identifies soybean trypsin inhibitor added to the samples of pancreatic juice to prevent autoactivation of pancreatic proteins during their analysis by the two-dimensional gel procedure. The upper abscissae give the isoelectric points of proteins; the left ordinate gives the apparent molecular weight $\times 10^{-3}$.

radioactive proteins injected into the blood circulation and those secreted into pancreatic juice collected between 2 and 7 h during the experiment. Although the radioactivity injected was exclusively represented by amylase (99.0%), only 30.7% of the radioactivity which subsequently appeared in pancreatic juice was represented by this enzyme. This contribution by amylase was identical to that observed in a crude zymogen granule fraction after labelling of rat pancreatic lobules with ^{35}S -methionine (29.9%). As the distribution of radioactivity after injection of ^{35}S -labelled amylase was indistinguishable from that resulting from injection of ^{35}S -methionine, we conclude that the amylase injected into the blood circulation was taken up by tissues in the rat and degraded to individual amino acids and that the resulting free ^{35}S -methionine, circulated through the

bloodstream, was used for synthesis of the entire complement of secretory proteins by the exocrine pancreas. No degradation was observed when ^{35}S -labelled amylase was incubated with rat serum at 37°C for 7 h. Presumably the bulk of amylase was taken up by the reticuloendothelial system because rat pancreatic amylase contains no attached carbohydrate. These findings indicate that rat pancreatic amylase is not directly transported from the blood circulation to the pancreatic juice. We therefore conclude that this component of the proposed enteropancreatic circulation for digestive enzymes does not exist.

Other studies strongly support our conclusions. Injection of ^{35}S -methionine-labelled guinea pig pancreatic proteins into the blood circulation of the rat resulted in the appearance of TCA-insoluble radioactivity in pancreatic juice after 1 h. Separation of proteins contained in pancreatic juice by the two-dimensional gel procedure, which discriminates guinea pig from rat exocrine pancreatic proteins, followed by fluorography, indicated that radioactivity was exclusively associated with rat proteins. The sensitivity of our analysis using the methods described would have allowed the detection of 0.002% of radioactive guinea pig pancreatic proteins.

In contrast with our studies analysing pancreatic juice, the analysis of bile after injection of pancreatic proteins into the blood circulation indicated that ^{35}S -labelled rat pancreatic proteins appeared rapidly, within 5 min, in biliary fluid. Figure 3 compares the Coomassie blue and fluorographic pattern of proteins contained in a sample of bile collected 20–30 min after injection of radioactive proteins into the blood circulation. Pancreatic proteins, which appeared in bile, represented 0.8% of the injected radioactivity and the appearance of negatively charged proteins was favoured over positively charged proteins (compare Figs 3*b* and 2*c*). These observations indicate not only the presence of a haematobiliary pathway for exocrine pancreatic proteins, but provide a powerful positive control for our studies into the possible existence of a haematopancreatic pathway. Taken together, our data convincingly demonstrate, to the level of 0.002%, that this latter putative pathway, at least in the rat, does not exist.

Received 5 January; accepted 27 May 1981.

1. Liebow, C. & Rothman, S. S. *Science* **189**, 472–474 (1975).
2. Götze, H. & Rothman, S. S. *Nature* **257**, 607–609 (1975).
3. Isenman, L. D. & Rothman, S. S. *Biochim. biophys. Acta* **585**, 321–332 (1979).
4. Beynon, R. J. & Kay, J. *Nature* **260**, 78–79 (1976).
5. Rothman, S. S., Götze, H. & Liebow, C. *Nature* **260**, 79 (1976).
6. Diamond, J. M. *Nature* **271**, 111–112 (1978).

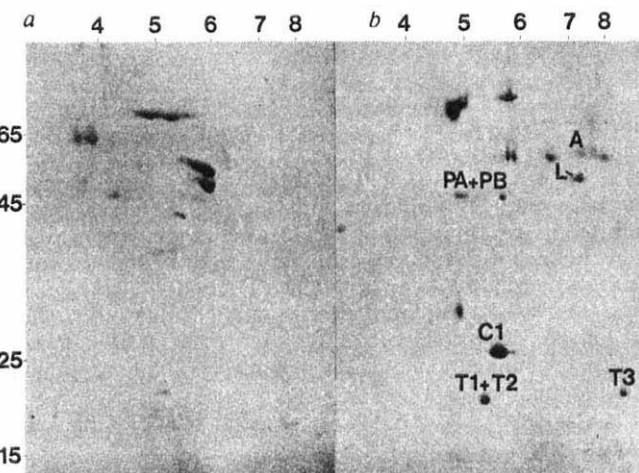


Fig. 3 Analysis of proteins contained in bile by two-dimensional isoelectric focusing/SDS-gel electrophoresis. Rat pancreatic proteins were labelled with ^{35}S -methionine as described in Fig. 1 legend. After injection of the entire complement of radioactive proteins into the blood circulation, bile samples were collected at 5-min intervals. *a* Shows the Coomassie blue-staining pattern and *b* the fluorographic pattern of proteins contained in the bile sample collected between 20 and 30 min. Radioactive pancreatic proteins observed in bile are labelled by abbreviations described in Fig. 2 legend. The radioactive protein with an isoelectric point of 4.8 and M_r (molecular weight) $\sim 72,000$ in *b* represents an acidic glycoprotein normally found in low levels in rat pancreatic secretion. The upper abscissae give the isoelectric points of proteins and the left ordinate gives the apparent molecular weight $\times 10^{-3}$.

7. Heinrich, H. C., Gabbe, E. E., Brüggmann, J. & Icagig, F. *Klin. Wschr.* **57**, 1295-97 (1979).
8. Lake-Bakaar, G., Rubio, C. E., McKavanagh, S., Potter, B. J. & Summerfield, J. A. *Clin. Sci. molec. Med.* **56**, 18 (1979); *Gut* **21**, 580-586 (1980).
9. Harper, A. & Scratcherd, T. in *The Exocrine Pancreas* (eds Howat, H. T. & Sarles, H.) 63 (Saunders, London, 1979).
10. Bieger, W. & Scheele, G. *Analyt. Biochem.* **109**, 222-230 (1980).
11. Scheele, G. *J. biol. Chem.* **250**, 5375-5385 (1975).
12. Scheele, G. A., Jacoby, R. & Carne, T. *J. Cell Biol.* **87**, 611-628 (1980).
13. Schramm, M. & Loyter, D. *Meth. Enzym.* **8**, 533-537 (1966).

Vasoconstriction with thromboxane A₂ induces ulceration of the gastric mucosa

B. J. R. Whittle, G. L. Kauffman & S. Moncada

Department of Prostaglandin Research, Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK

Thromboxane A₂ (TXA₂) is formed from the fatty acid precursor, arachidonic acid (AA), by the sequential actions of the enzymes cyclooxygenase and thromboxane synthetase¹. It is readily generated by platelets and is a potent vasoconstrictor and inducer of platelet aggregation^{1,2}. Studies in dogs show that the addition of AA to blood flowing through an incubation coil leads to production of TXA₂, characterized by its instability and its ability to contract isolated strips of vascular tissue³. We now report that the formation of TXA₂ from AA in blood can be inhibited by 1-benzylimidazole, a thromboxane synthetase inhibitor, and selectively by very low doses of indomethacin. In conditions of TXA₂ inhibition, a vasodilator AA metabolite, presumably prostacyclin, is formed in the stomach. We show that TXA₂ is a powerful gastric vasoconstrictor and can extensively damage the gastric mucosa, and suggest that it may be involved in the pathogenesis of certain ulcerative disorders of the stomach.

Segments of acid-secreting fundic mucosa from the stomachs of pentobarbitone-anaesthetized dogs were encased in a lucite chamber as described by others⁴. The branches of the splenic artery and vein, which supplied the mucosal flap, were isolated, and all other vascular connections supplying the greater curvature of the stomach and the spleen were ligated. Each dog was heparinized (500 IU per kg), and the splenic artery was cannulated close to the mucosal flap and perfused (10 ml min⁻¹) with blood from a cannulated femoral artery with a constant flow roller-pump (Watson Marlow). Changes in perfusion pressure, indicating vascular resistance, were recorded with a transducer connected to the arterial blood line close to the stomach. Resting perfusion pressure was 80-100 mm Hg. Injection ports were sited in the arterial blood line, either close to the stomach (with a 3-s delay before reaching the stomach) or distally into a delay coil (allowing 30-s incubation in blood before the injected substances reached the stomach). Mean arterial blood pressure was measured via a cannula in a femoral artery and drugs could be administered intravenously (i.v.) through a cannula in a femoral vein.

Bolus injection of AA (25-200 µg intra-arterially (i.a.)) into the 30-s incubation coil produced a dose-dependent rise in gastric perfusion pressure, with 100 µg AA producing a near-maximal increase of 56 ± 9 mm Hg (mean ± s.e.m. mean; 13 animals). This vasoconstriction was accompanied by localized blanching in the superficial mucosa and increased motility of the tissue within the chamber.

We investigated whether the increase in perfusion pressure induced by AA could be due to platelet aggregates blocking the microcirculation. Although some workers have shown that dog platelets do not aggregate when they form TXA₂ from AA⁵,

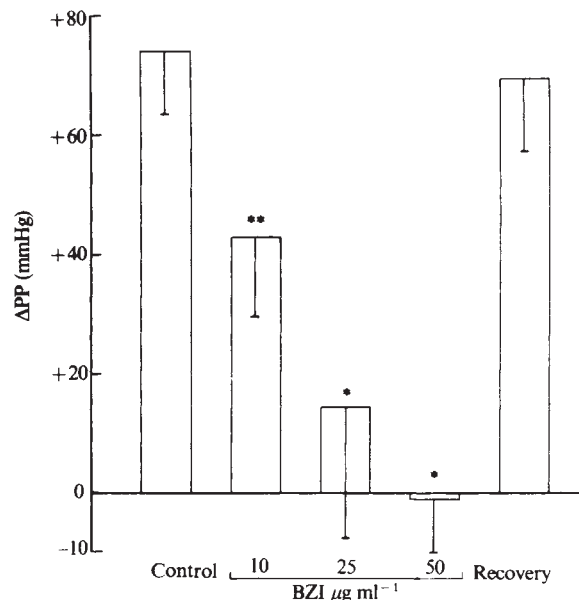


Fig. 1 Vasoconstriction in the canine gastric circulation by TXA₂ generated from arachidonic acid in blood, and its inhibition by 1-benzylimidazole in five dogs. AA (Grade I; Sigma) was stored in *n*-hexane (10 mg ml⁻¹) at -20 °C. After evaporation of the *n*-hexane with N₂ from 1-ml aliquots, sodium hydroxide (0.5 M in methanol, 0.5 ml) was added and subsequently evaporated to dryness. The residue was dissolved (10 mg ml⁻¹) in 1 M Tris buffer (pH 10.5, 4 °C), and stored in a light-proof container on ice. 'Control' represents response to AA (100 µg in 10-µl volume) injected i.a. into a delay coil so as to incubate with blood for 30 s before reaching the stomach. BZI (10-50 µg ml⁻¹ final concentration in blood) was infused i.a. (0.1 ml min⁻¹) so as to incubate with blood 30 s before reaching injection site for AA. Results, shown as change in gastric perfusion pressure (ΔPP), are mean ± s.e.m. of five experiments, where level of statistical difference from control (using paired data *t*-test) is shown as: *, *P* < 0.05; **, *P* < 0.01.

infusions of very high doses of AA i.a. to the canine stomach, which elevate perfusion pressure⁷, may lead to thrombotic occlusion of the gastro-epiploic artery⁶. We have determined gastric arterio-venous (AV) differences in platelet counts in plasma after AA administration. Samples (2 ml) of blood collected in trisodium citrate (0.32%) were rapidly centrifuged in a modified Eppendorf bench centrifuge⁸ and the platelet count in the plasma determined in a Coulter counter. In four dogs, the AV platelet count after i.a. administration of 100 µg AA into the delay coil was not significantly (*P* > 0.05) different from that obtained without AA in control conditions, indicating that platelet aggregation with microemboli plugging of the gastric microcirculation is not a major factor in this vasoconstriction.

1-Benzylimidazole (BZI), like imidazole², is a selective inhibitor of TXA₂ formation in both human⁹ and dog¹⁰ platelets. To investigate its action, the hydrogen fumarate salt of BZI was dissolved in isotonic saline and infused into the delay coil so that it mixed with the blood for 30 s before reaching the AA injection site (thus incubating for 1 min in blood before reaching the stomach). In five dogs, BZI infusion (10-50 µg ml⁻¹, final concentration in the blood) alone had no effect on resting perfusion pressure in the gastric circulation. At 3 min after starting the infusion of BZI (10-50 µg ml⁻¹), the vasoconstrictor response to AA (100 µg into the delay coil) was dose-dependently reduced, presumably reflecting inhibition of TXA₂ formation (Fig. 1). The BZI concentration reducing the vasoconstrictor response by 50% (13 µg ml⁻¹; 50 µM) is similar to that described for the inhibition of thromboxane formation from

Table 1 Effect of topical application for 1 h of acid (100 mM HCl) or acid (100 mM) plus sodium taurocholate (5 mM) either alone or during TXA₂-induced gastric vasoconstriction

	Acid	Acid + AA	Acid + Tauro	Acid + Tauro + AA
p.d. (mV)	-60 ± 3	-57 ± 6	-44 ± 3*	-43 ± 2**
ΔH ⁺ (μmol)	-234 ± 202	-8 ± 239	-837 ± 224*	-403 ± 240
Lesion area (mm ²)	0	0	25 ± 11*	278 ± 55**
n	4	3	3	4

AA was infused i.a. (5 μg ml⁻¹ or 50 μg min⁻¹) into a 30-s incubation coil to generate TXA₂ for the latter 30 min of the study period. Results, which show the p.d., acid back-diffusion (ΔH⁺) and lesion area of gastric mucosa over 1 h, are mean ± s.e.m. of values from *n* dogs. Significant differences from acid-alone group is shown by *, *P* < 0.05; **, *P* < 0.01.

¹⁴C-AA by canine platelets *in vitro*¹⁰. Infusion of BZI into the injection port close to the stomach failed to prevent the vasoconstrictor response to AA, indicating that the generation of TXA₂ was primarily in the blood and not in the stomach (Fig. 2). With 50 μg ml⁻¹ of BZI in the delay coil, a net vasodilator response to AA was observed in three of five experiments, suggesting that reduced TXA₂ formation in blood can unmask the biotransformation of AA to a metabolite, presumably prostacyclin¹¹, which is a vasodilator in the stomach (Fig. 2). Indeed, even in the absence of BZI, close arterial administration of AA (100 μg), with only a 3-s delay before reaching the stomach, significantly (*P* < 0.01) reduced perfusion pressure (-23 ± 8 mm Hg, *n* = 8).

Within 3 min of terminating the BZI infusion to the incubation coil, the vasoconstrictor response to AA had returned to control values (Figs 1, 2). This system is thus useful for studying locally acting drugs without treatment of the whole animal.

The effect of a systemically administered cyclooxygenase inhibitor¹² was also studied in six dogs. A low dose of indomethacin (25 μg per kg intravenously (i.v.)), given 5 min before injection of AA into the 30-s delay coil, significantly (*P* < 0.05) reduced the vasoconstrictor response (Fig. 3). With higher doses of indomethacin (50–500 μg per kg, i.v.) the vasoconstrictor response was converted into a vasodilator response, whereas with 1–5 mg per kg, i.v., indomethacin abolished all the vasoactive effects of AA. We conclude that low doses of indomethacin selectively inhibit the cyclooxygenase in platelets, thus reducing the formation of the vasoconstrictor, TXA₂, while cyclooxygenase in the stomach retains its ability to transform AA into the vasodilator metabolite, prostacyclin. At higher doses, the cyclooxygenase in the stomach also becomes inhibited, with subsequent abolition of the vasodilator response. Such differential actions may explain the small vasodilation in

the canine gastric circulation with infusion of high doses of AA after indomethacin pretreatment⁷.

As discussed previously for aspirin¹³, our present results in the dog suggest that indomethacin can selectively inhibit platelet cyclooxygenase *in vivo*. Whether such differential sensitivity would be of clinical value for the inhibition of platelet aggregation must await further studies as the more chronic administration of low doses of indomethacin may inhibit cyclooxygenase in most tissues. Previous studies with indomethacin at anti-inflammatory doses in rats did not reveal any selectivity between inhibitory effects on prostaglandin production in the inflammatory exudate and in the gastric mucosa (unlike sodium salicylate and the anti-inflammatory compound, BW755C) while the

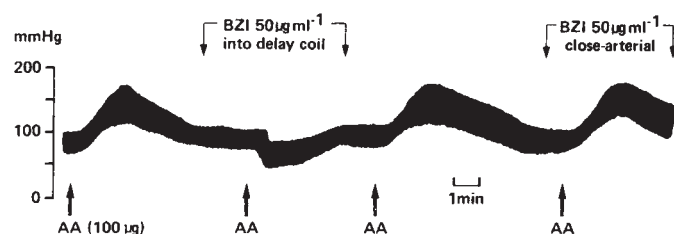


Fig. 2 Vasoconstriction in the canine gastric circulation induced by TXA₂, generated from AA (100 μg in 10 μl volume) injected into a 30-s delay coil. In this experiment, the vasoconstrictor response was converted to a vasodilator response when 1-benzylimidazole (BZI, 50 μg ml⁻¹ final concentration in blood) was infused (0.1 ml min⁻¹) into the delay coil. BZI, infused locally to the stomach, failed to prevent the vasoconstrictor response to AA, indicating that TXA₂ was being generated in blood and not in the stomach.

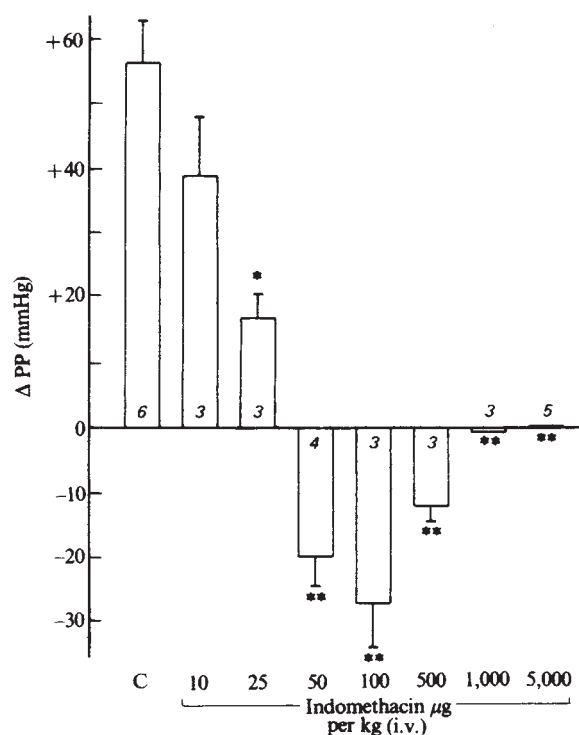


Fig. 3 Effect of indomethacin (10–5,000 μg per kg) administered i.v. on the gastric vasoactive actions of AA (100 μg; injected i.a. into a delay coil, so as to incubate with blood for 30 s before reaching the stomach). Indomethacin was dissolved in 5% w/v sodium bicarbonate solution, and the stock (20 mg ml⁻¹) diluted to desired concentration. Indomethacin in progressively increasing doses reduced the control (C) vasoconstriction with TXA₂ generated from AA, and converted it into a vasodilator response, and subsequently abolished all vasoactive responses. Results, shown as increase or decrease in gastric perfusion pressure (ΔPP), are the mean ± s.e.m. Number of animals used is shown in each bar. *, *P* < 0.02; **, *P* < 0.001.

inhibition of gastric cyclooxygenase was correlated with the production of gastric erosions¹⁴.

As mucosal ischaemia may be involved in the pathogenesis of gastric ulceration, we investigated whether the gastric vasoconstrictor actions of TXA₂ could induce or enhance gastric damage. Although vasoconstriction induced in dogs by vasopressin alone does not induce marked erosion, it does potentiate the damage from topical application of a low concentration of bile salts⁴. We have therefore compared the gastric lesions after 1 h topical application of acid (100 mM HCl) or acid plus the bile salt (100 mM HCl, 5 mM taurocholate) both alone, and during generation of the TXA₂ by infusion of AA (50 µg min⁻¹; 5 µg ml⁻¹ final concentration in blood) into the 30-s delay coil. This infusion produced a sustained increase in gastric perfusion pressure of 56 ± 7 mm Hg (*n* = 10), and is comparable with the AA concentration which generated TXA₂ in dog blood, as detected by bioassay on vascular tissues³. In a randomized experiment in eight dogs, the acid back-diffusion (measured by titration) and electrical potential difference (p.d.) across the mucosa, indices of mucosal permeability¹⁵, were also determined in a two-compartment chamber.

Topical application of acid plus taurocholate for 1 h significantly increased acid loss and lowered p.d., but caused only slight gastric mucosal damage (measured in terms of lesion area with calipers) as shown in Table 1. Vasoconstriction induced by TXA₂, generated from AA in blood, did not significantly alter acid back-diffusion or p.d. during application of topical acid, nor lead to observable damage (Table 1). However, in the presence of acid plus taurocholate, substantial damage of the mucosa was observed during vasoconstriction with TXA₂. Necrosis and punctate bleeding became apparent within 10 min and large areas of the mucosal surface became damaged within 30 min (Table 1). After terminating the AA infusion, frank bleeding from the eroded surface was observed.

This work using TXA₂ confirms that vasoconstrictor agents can readily induce ulceration of the gastric mucosa in conditions where the mucosal surface had been in contact with low concentrations of bile salts^{4,16}. Although back-diffusion of acid alone into the mucosa seems to cause only slight gastric damage, concurrent vasoconstriction allows the accumulation of these back-diffusing hydrogen ions, leading to mucosal tissue acidification and thus necrosis. Our findings that the TXA₂ generated from AA in blood is a powerful gastric vasoconstrictor and leads to extensive mucosal damage may implicate endogenous TXA₂ in the pathogenesis of gastric ulceration. Clinical conditions such as stress or shock might lead to local platelet activation with subsequent TXA₂ generation in the stomach, or lead to an imbalance between prostacyclin-TXA₂ levels, and hence induce gastric mucosal damage.

We thank Dr P. Thorogood for preparing 1-benzylimidazole and Dr J. R. Vane for helpful discussions.

Received 21 March; accepted 10 June 1981.

1. Hamberg, M., Svensson, J. & Samuelsson, B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2994–2998 (1975).
2. Moncada, S. *et al. Prostaglandins* **13**, 611–618 (1977).
3. Mullane, K. M., Dusting, G. J., Salmon, J. A., Moncada, S. & Vane, J. R. *Eur. J. Pharmac.* **54**, 217–218 (1979).
4. Ritchie, W. P. *Gastroenterology* **68**, 699–707 (1975).
5. Chignard, M. & Vargaftig, B. B. *Eur. J. Pharmac.* **38**, 7–18 (1976).
6. Konturek, S. J., Mikos, E., Pawlik, W. & Walus, K. M. *J. Physiol., Lond.* **286**, 15–28 (1979).
7. Walus, K. M., Gustaw, P. & Konturek, S. J. *Prostaglandins* **20**, 1089–1102 (1980).
8. Whittle, B. J. R., Moncada, S., Whiting, F. & Vane, J. R. *Prostaglandins* **19**, 605–627 (1980).
9. Tai, H. H. & Yuan, B. *Biochem. biophys. Acta* **539**, 162–172 (1978).
10. Harris, R. H. *et al. Adv. Prostaglandin Thromboxane Res.* **6**, 457–461 (1980).
11. Moncada, S., Gryglewski, R. J., Bunting, S. & Vane, J. R. *Nature* **263**, 663–665 (1976).
12. Vane, J. R. *Nature new Biol.* **231**, 232–235 (1971).
13. Moncada, S. & Vane, J. R. *New Engl. J. Med.* **300**, 1142–1147 (1979).
14. Whittle, B. J. R., Higgs, G. A., Eakins, K. E., Moncada, S. & Vane, J. R. *Nature* **284**, 271–273 (1980).
15. Davenport, H. W. *Gastroenterology* **46**, 245–253 (1964).
16. Whittle, B. J. R. *Br. J. Pharmac.* **60**, 455–460 (1977).

Mobility of polypeptide chain in the pyruvate dehydrogenase complex revealed by proton NMR

Richard N. Perham & Harry W. Duckworth*

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

Gordon C. K. Roberts

National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Recent studies of several small proteins by NMR spectroscopy and X-ray crystallography have clearly demonstrated significant internal mobility in their structures (see, for example, refs 1–9), which can involve not only amino acid side chains but also larger regions of polypeptide chain. Occasionally a plausible function for this mobility has been suggested^{1,9}, but there has been no conclusive evidence for a direct connection between intramolecular mobility and a defined step in an enzymatic mechanism. The pyruvate dehydrogenase (PDH) multienzyme complex of *Escherichia coli* (molecular weight (*M*_r) 4.5–6 × 10⁶) is one of the largest well defined assemblies of proteins known, comprising multiple copies of three different enzymes^{10,11}. The substrate is carried in thioester linkage by lipoyl-lysine residues of the lipoate acetyltransferase component, the structural core of the complex. The lipoyl-lysine residues act as swinging arms, carrying substrate between the catalytic centres of the three enzymes^{12–15} and between lipoic acid residues attached to different subunits in the lipoate acetyltransferase core^{16–18}. It has been conjectured that the lipoic acid-containing regions of polypeptide chain might be flexible^{19,20} and therefore able to increase greatly the effective radius of a swinging arm¹⁹. We report here unexpectedly sharp lines in the 270-MHz proton NMR spectrum of the enzyme complex that are attributed to remarkable conformational mobility of large regions of polypeptide chain carrying the lipoic acid residues. This mobility would enhance the functional connection of active sites in a multisubunit structure.

A schematic mechanism of the PDH complex is shown in Fig. 1. A 270-MHz proton NMR spectrum of the intact complex is shown in Fig. 2a, the most striking feature of which is the appearance of several sharp resonance lines (between –0.6 and –3 p.p.m. from dioxan) attributable to the protein. The complex has an *M*_r of ~6 × 10⁶ (ref. 11), a diameter of at least 30 nm (refs 10, 11) and a rotational correlation time of ~10⁻⁵ s (ref. 14). A methylene proton moving with this correlation time would have a linewidth of ~8.5 kHz, and indeed there is an extremely broad envelope of resonance underlying the sharp lines in Fig. 2a. However, resonances with linewidths of only ~50 Hz must mean that some amino acid residues have substantial freedom to move rapidly with respect to the enzyme complex.

The most prominent sharp resonance in the spectrum is at –2.34 p.p.m., close to the positions of the methyl resonances of alanine (–2.32 p.p.m.) and threonine (–2.48 p.p.m.) in simple peptides. There is also a clear shoulder at –2.77 p.p.m., which is similar to the chemical shifts of the methyl protons of valine, leucine and isoleucine (–2.82 to –2.74 p.p.m.). Methyl groups in proteins are commonly found to rotate very rapidly about the carbon–carbon bond attaching them to the molecule^{1,2}. However, this motion is not sufficient to explain the observed linewidths for these resonances. Rapid rotation of a methyl

* Permanent address: Department of Chemistry, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2.

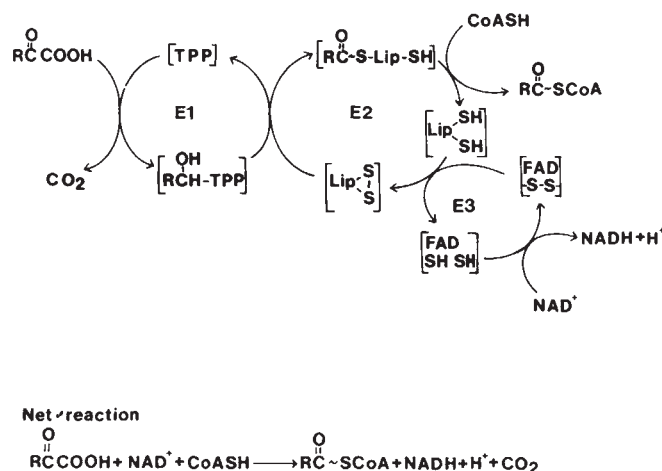


Fig. 1 The reaction mechanism of the PDH multienzyme complex. E1 (pyruvate decarboxylase, EC 1.2.4.1), E2 (lipoate acetyltransferase, EC 2.3.1.12) and E3 (lipoamide dehydrogenase, EC 1.6.4.3) represent the three enzymatic activities. TPP, thiamine pyrophosphate; R = CH₃—.

group about this single axis can decrease the linewidth by no more than a factor of 4 (refs 21, 22) to give a predicted methyl linewidth of ~4 kHz in the present case—almost two orders of magnitude greater than that observed. It is clear that motion about several bonds must be involved, and that the sharp signals in Fig. 2a must arise from a region of the protein where the backbone, as well as the side chains, has substantial mobility.

It is impossible to integrate these spectra accurately owing to the obvious difficulties of drawing an appropriate baseline, but approximate integration of spectra obtained with a 5-s pulse interval indicates that several hundred protons per enzyme protomer contribute to the sharp lines in the spectrum. We define a protomer as 1/24th part of the enzyme complex, as the particle appears to be a 24-fold repeat of this structure, based on an E2 core of 24 polypeptide chains arranged with octahedral symmetry^{10,11}. Thus more than a few amino acids are involved and substantial regions of polypeptide chain must be mobile. The observed intensity ratio of the (Ala+Thr) to the (Val+Leu+Ile) resonances is greater than would be expected from the amino acid composition of the complex (see refs listed in ref. 23), and in addition no sharp resonances from aromatic protons are observed. Clearly we are not observing a general mobility of all the polypeptide chains in the complex, but rather a high degree of mobility for specific regions of the primary structure.

No change in the NMR spectrum was observed when the complex was converted to the acyl enzyme by treatment with pyruvate (0.11 mM), thiamine pyrophosphate (0.11 mM) and MgCl₂ (5.7 mM) in the absence of coenzyme A and NAD⁺ (see mechanism in Fig. 1). Likewise, no change was observed when the lipoic acid residues were reduced by treatment with NADH (0.15 mM) and NAD⁺ (0.05 mM) or when this reduced enzyme complex was acetylated by treatment with acetyl CoA (0.14 mM).

The PDH complex can be resolved into two subcomplexes: an E1–E2 subcomplex by selective removal of the E3 subunits in 4 M urea, and an E2–E3 subcomplex by selective removal of the E1 subunits at pH 10 (refs 24, 25). Both subcomplexes have a large *M_r* because they are based on the octahedral (24-chain) core of the E2 component. Their appearance on SDS-gel electrophoresis is shown in Fig. 3 and their proton NMR spectra in Fig. 2. The most prominent features of the spectrum of the intact complex are retained in the spectra of the subcomplexes. As the only type of polypeptide chain common to all three structures is that of lipoate acetyltransferase (E2), we infer that the most mobile regions of the intact PDH complex belong to this core enzyme.

The isolated E1 and E3 components are dimers, with *M_r*s of ~200,000 and 112,000 respectively¹⁰. Their proton NMR spectra (Fig. 4) differ from those of the intact complex and subcomplexes (Fig. 2), lacking the prominent (Ala+Thr) resonance at -2.34 p.p.m., but having clearly detectable resonances from aromatic protons. The lines in these spectra have apparent widths (~100 Hz) similar to those expected for molecules of this size^{1,26,27}, and there is no indication of substantial internal motion. These experiments, together with those on the subcomplexes, clearly implicate the E2 component as the source of the sharp resonances in the spectrum of the complex. Attempts to obtain a spectrum of the isolated E2 component were frustrated by its tendency to aggregate and precipitate at the high protein concentrations required.

Each E2 chain of the *E. coli* complex contains two lipoic acid residues which become reductively acetylated by substrate (see Fig. 1) during the enzymatic reaction^{16,17,28}. Limited trypsin treatment of the complex releases fragments of the E2 polypeptide chain carrying these lipoyl groups, whereas E1 and E3 components remain bound to the residual part of the E2 core in an inactive complex which still has a very high *M_r*^{19,20,29}. This limited proteolysis suggested that the lipoic acid-containing

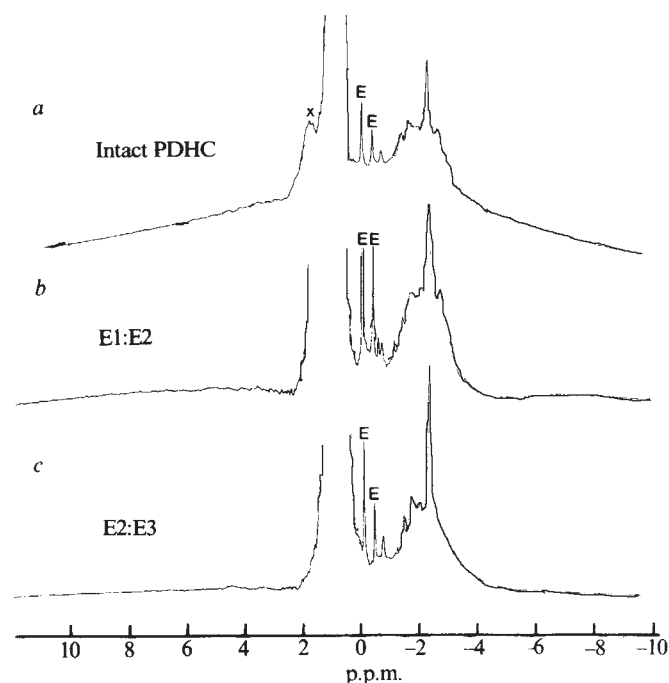


Fig. 2 The proton NMR spectra of *E. coli* PDH complex and its subcomplexes. Complexes¹⁶ were pelleted in the ultracentrifuge (3 h, 45,000g), dissolved in D₂O buffer, apparent *pD* 7.6, containing 20 mM potassium phosphate and 2 mM sodium EDTA, and dialysed exhaustively against the same buffer. Spectra were obtained at 270 MHz using a Bruker WH270 spectrometer. Quadrature phase detection was used, with a 6-kHz spectral width and 0.34-s pulse intervals. The free-induction decay was recorded in 4,096 data points and 2,000 decays were averaged. Before Fourier transformation, the free-induction decay was multiplied by an exponential corresponding to a line broadening of 5 Hz and the data table was extended to 8,192 points with zeros. Chemical shifts are expressed relative to internal dioxan. The peaks marked E arise from EDTA. *a*, Intact complex, 70 mg ml⁻¹; *s*_{20,w} (sedimentation coefficient corrected to water at 20 °C) of a portion of this sample, diluted to 3.0 mg ml⁻¹, was 58.7 S. The feature marked x is an instrumental artefact. *b*, E1–E2 subcomplex, prepared by resolution of intact complex on hydroxyapatite in 6 M urea, a method similar to that in ref. 24; 29 mg ml⁻¹. This preparation was a cloudy suspension which was not examined by analytical ultracentrifugation. *c*, E2–E3 subcomplex, prepared by treatment of intact complex at pH 10 according to refs 24,25; 57 mg ml⁻¹; *s*_{20,w} at 3.0 mg ml⁻¹ = 30.9 S.

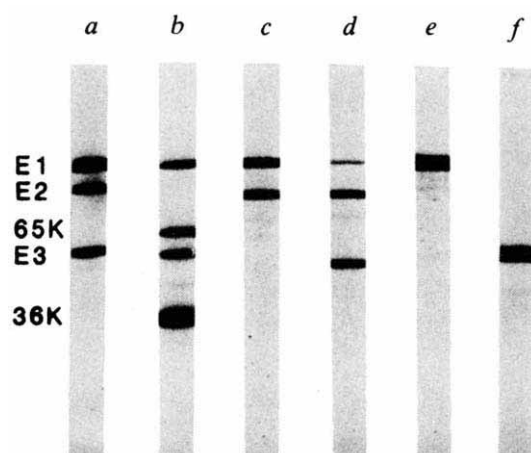


Fig. 3 Analysis by SDS-polyacrylamide gel electrophoresis of *E. coli* PDH complexes and subcomplexes whose NMR spectra are shown in Figs 2 and 4. All samples were run on standard 7.5% gels in phosphate buffer¹⁶. *a*, Intact enzyme complex; *b*, enzyme complex after limited digestion with trypsin and gel filtration on Sepharose 6B-CL¹⁹; *c*, E1-E2 subcomplex prepared by removal of E3 subunits in the presence of 6M urea²⁴; *d*, E2-E3 subcomplex prepared by removal of E1 subunits at pH 10 (refs 24, 25); *e*, free E1 (dimers); *f*, free E3 (dimers). Fragments with apparent M_r of 65,000 and 36,000 in *b* are tryptic products of the E1 and E2 chains, respectively²⁹.

regions of the lipoyl acetyltransferase core protrude physically in the complex^{19,20,29}, and may form a 'fuzz' of lipoyl domains surrounding the E2 core in favourably oriented electron micrographs²⁰. The largest identifiable fragment containing lipoyl acid residues has an apparent M_r of 45,000²⁰ but this is rapidly broken down further by trypsin. Some degradation of the E1 chains from an apparent M_r of 100,000 to 65,000 is usually found and the E2 core is then composed of fragments of the E2 chain with an apparent M_r of 36,000, compared with an apparent M_r of 80,000 for the intact E2 polypeptide^{20,29} (see Fig. 3). Figure 4*d* shows the proton NMR spectrum of a trypsin-treated complex, gel-filtered on Sepharose 6B to remove the lipoyl acid-containing peptides¹⁹. All the sharpest peaks in the spectrum of the intact complex have disappeared, indicating that the proteolysis has removed the bulk of the highly mobile region or regions of the E2 chain. The spectrum of the tryptic core still contains resonances in the aliphatic region which are too narrow to arise from a rigid complex of this size, although they are much broader than the sharp components of the spectrum of the intact complex. They indicate the presence of regions of moderate internal mobility in the tryptic core.

The almost unfettered motion of the lipoyl groups revealed by ESR spectroscopy^{14,15} is accompanied by much conformational mobility of polypeptide regions of the lipoyl acetyltransferase core that encompass the lipoyl-lysine residues. These are intrinsic properties of the enzyme protein and must not be confused with substrate-induced conformational changes. The chemical shifts of the sharp resonances in the NMR spectra are consistent with a random coil-like structure for the mobile regions, although some parts of these regions may have defined three-dimensional structures. Such structured regions near the lipoyl acid residues might facilitate recognition by the enzyme active sites through protein-protein interaction.

Conformational mobility may explain some puzzling features of the enzyme complex. It might help to span the gaps between E1, E2 and E3 active sites¹⁹, which may be larger than a single lipoyl-lysine swinging arm could cover in a rigid protein³⁰⁻³², and it is presumably important in active-site coupling¹⁶⁻¹⁸. Several lines of evidence indicate that up to about half of the

lipoyl acid residues of the *E. coli* PDH complex can be removed by limited proteolysis with trypsin²⁰ or inactivated by treatment with *N*-ethylmaleimide³³ without much loss of overall catalytic activity. The overall catalytic activity of the complex is directly proportional to its E1 content^{16,18,34} and the rate-limiting step in the mechanism must involve E1^{18,35}. Loss of catalytic activity would lag behind the loss or modification of lipoyl acid residues if one lipoyl acid residue could take over the role of another^{20,33,36,37}. Conformational mobility in the E2 core might facilitate this by allowing a given E1 active site to be visited by more than one lipoyl acid residue, even a lipoyl acid residue bound to a different E2 chain^{16,33,36}. The dihydrolipoyl group could in turn have a limited choice of E3 active sites at which to be reoxidized³⁸. That this mobility is vital in the functioning of the complex is reinforced by the observation (H.W.D., R. Jaenicke, R.N.P., G.C.K.R. and E. Wawrzyniak, unpublished)

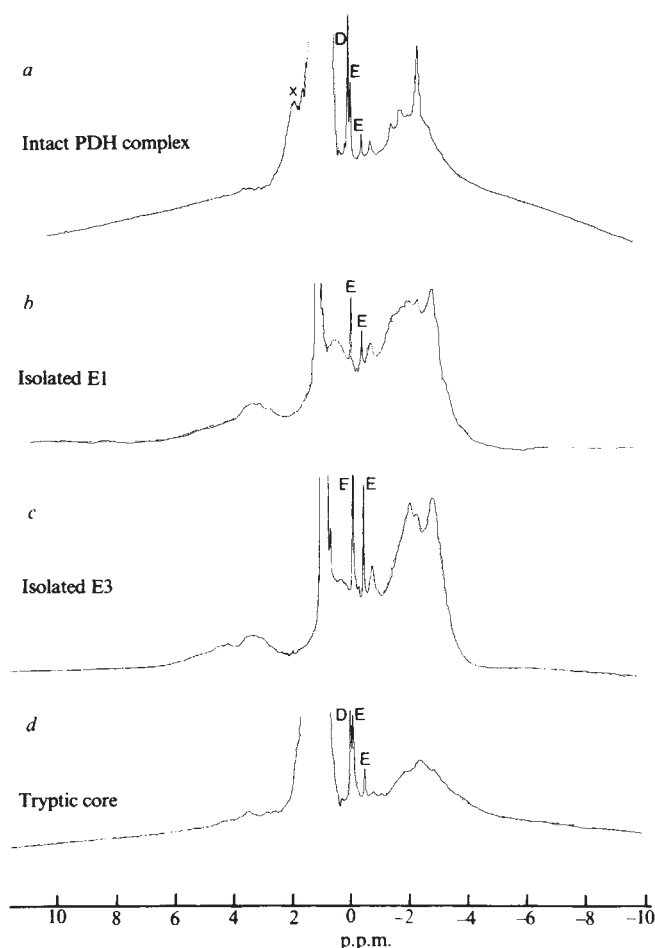


Fig. 4 The proton NMR spectra of some derivatives of *E. coli* PDH complex. Protein samples were concentrated by ultracentrifugation (*a*, *d*) or ultrafiltration (*b*, *c*), and were then dialysed against the D₂O buffer and examined by NMR spectroscopy as described in Fig. 2 legend. The signal marked D in *a* and *d* is due to dioxan, ~1 mM, added as a reference; peaks E arise from EDTA; x is an instrumental artefact. *a*, Intact complex from Fig. 1. *b*, E1 component, obtained by the resolution procedure that yielded E2-E3 subcomplex (Fig. 2*c*); 51 mg ml⁻¹; $s_{20,w}$ at 3.0 mg ml⁻¹ = 9.65 S, as expected for the dimer, M_r = 200,000 (ref. 10). *c*, E3 component, obtained by the resolution procedure that yielded E1-E2 subcomplex (Fig. 2*b*); 12 mg ml⁻¹; $s_{20,w}$ at 3.0 mg ml⁻¹ = 6.03 S, as expected for the dimer, M_r = 112,000 (ref. 10). *d*, Core complex obtained by limited proteolysis of native PDH complex with trypsin and gel filtration on Sepharose 6B-CL, as in ref. 19; 66 mg ml⁻¹; $s_{20,w}$ at 3.0 mg ml⁻¹ = 41.2 S.

that the PDH complexes of *Bacillus stearothermophilus* and ox heart and the 2-oxoglutarate dehydrogenase complexes of *E. coli* and ox heart all display very similar NMR spectra, indicating similar conformational mobility.

This work was supported in part by a research grant from the SRC (to R.N.P.). H.W.D. is a Travelling Fellow of the Nuffield Foundation. We thank Miss M. B. Thomas for technical assistance and D. Reed for performing the ultracentrifuge runs.

Received 17 February; accepted 2 June 1981.

- Gurd, F. R. N. & Rothgeb, T. M. *Adv. Protein Chem.* **33**, 73–165 (1979).
- Williams, R. J. P. *Biol. Rev.* **54**, 389–437 (1979).
- Frauenfelder, H., Petsko, G. A. & Tsemoglou, D. *Nature* **280**, 558–563 (1979).
- Artymuik, P. I. *et al. Nature* **280**, 563–568 (1979).
- Campbell, I. D., Dobson, C. M., Moore, G. R., Perkins, S. J. & Williams, R. J. P. *FEBS Lett.* **70**, 96–100 (1976).
- Wagner, G., De Marco, A. & Wüthrich, K. *Biophys. Struct. Mech.* **2**, 139–158 (1976).
- Ribeiro, A. A., King, R., Restivo, C. & Jardetzky, O. *J. Am. chem. Soc.* **102**, 4040–4051 (1980).
- Jardetzky, O., Adasaka, K., Vogel, D., Morris, S. & Holmes, K. C. *Nature* **273**, 564–566 (1978).
- Highsmith, S. *et al. Biochemistry* **18**, 4238–4244.
- Reed, L. J. *Acc. chem. Res.* **7**, 40–46 (1974).
- Danson, M. J. *et al. J. molec. Biol.* **129**, 603–617 (1979).
- Green, D. E. & Oda, T. *J. Biochem., Tokyo* **49**, 742–757 (1961).

- Koike, M., Reed, L. J. & Carroll, W. R. *J. biol. Chem.* **238**, 30–39 (1963).
- Ambrose, M. C. & Perham, R. N. *Biochem. J.* **159**, 429–432 (1976).
- Grande, H. J., Van Telgen, H. J. & Veeger, C. *Eur. J. Biochem.* **71**, 509–518 (1976).
- Bates, D. L., Danson, M. J., Hale, G., Hooper, E. A. & Perham, R. N. *Nature* **268**, 313–316 (1977).
- Collins, J. H. & Reed, L. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4223–4227 (1977).
- Danson, M. J., Fersht, A. R. & Perham, R. N. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5386–5390 (1978).
- Hale, G. & Perham, R. N. *FEBS Lett.* **105**, 263–266 (1979).
- Bleile, D. M., Munk, P., Oliver, R. M. & Reed, L. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4385–4389 (1979).
- Woessner, D. E., Snowden, B. S. Jr & Meyer, G. H. *J. chem. Phys.* **50**, 719–721 (1969).
- Werbelow, L. G. & Marshall, A. G. *J. Am. chem. Soc.* **95**, 5132–5134 (1973).
- Bates, D. L., Harrison, R. A. & Perham, R. N. *FEBS Lett.* **60**, 427–430 (1975).
- Reed, L. J. & Willms, C. R. *Meth. Enzym.* **9**, 247–265 (1966).
- Coggins, J. R., Hooper, E. A. & Perham, R. N. *Biochemistry* **15**, 2527–2533 (1976).
- Wüthrich, K. *NMR in Biological Research* (Elsevier, Amsterdam, 1976).
- Jardetzky, O. & Roberts, G. C. K. *NMR in Molecular Biology* (Academic, New York, 1981).
- Danson, M. J. & Perham, R. N. *Biochem. J.* **159**, 677–682 (1976).
- Hale, G. & Perham, R. N. *Eur. J. Biochem.* **94**, 119–126 (1979).
- Shepherd, G. B. & Hammes, G. G. *Biochemistry* **15**, 311–317 (1976).
- Angelides, K. J. & Hammes, G. G. *Biochemistry* **18**, 1223–1229 (1979).
- Scouten, W. H. *et al. Eur. J. Biochem.* **112**, 9–16 (1980).
- Ambrose-Griffin, M. C., Danson, M. J., Griffin, W. G., Hale, G. & Perham, R. N. *Biochem. J.* **187**, 393–401 (1980).
- Angelides, K. J. & Hammes, G. G. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4877–4880 (1978).
- Akiyama, S. K. & Hammes, G. G. *Biochemistry* **19**, 4208–4213 (1980).
- Cate, R. L., Roche, T. E. & Davis, L. C. *J. biol. Chem.* **255**, 7556–7562 (1980).
- Berman, J. N., Chen, G.-X., Hale, G. & Perham, R. N. *Biochem. J.* (submitted).
- Reed, L. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 3068–3072 (1975).

Delayed recruitment of maternal histone H3 mRNA in sea urchin embryos

Dan E. Wells, Richard M. Showman, William H. Klein & Rudolf A. Raff

Program in Molecular, Cellular and Developmental Biology, Department of Biology, Indiana University, Bloomington, Indiana 47405, USA

During the early stages of embryogenesis, the sea urchin embryo uses maternally synthesized mRNA stored in the egg as inactive messenger ribonucleoproteins (mRNPs)¹. Release of this mRNA for translation allows the embryo to develop even in the absence of new mRNA synthesis. Comparison using two-dimensional gel electrophoresis shows that the same spectrum of prevalent proteins are synthesized by eggs and zygotes 30–60 min after fertilization². Thus, unlike the case of the clam *Spisula* in which there are several prominent changes in translation associated with fertilization³, sea urchins have been thought to show little or no regulation of translation of specific mRNA sequences. Histone mRNAs are major components of the mRNA pool, comprising as much as 4–8% of the total mRNA of eggs⁴, yet their products, histones, would not normally be detectable on two-dimensional gels. We report here that mRNA complementary to a histone H3 cloned probe remains in the inactive mRNA pool for 90 min after fertilization before it begins to be translated. This is long after the rapid increase in overall protein synthesis, using stored mRNAs, has begun. As we can demonstrate no significant synthesis of histone H3 mRNA during this period, stored histone H3 mRNA must be subject to sequence-specific translational regulation.

Patterns of protein synthesis during early sea urchin embryogenesis have supported the idea of a rapid linear release of stored maternal mRNA from the mRNA pool. The kinetics of protein synthesis require that the release of stored mRNA for translation starts 2 min after fertilization⁵. This rapid initiation of the release of stored mRNAs is further indicated by the kinetics of recruitment of free ribosomes into polysomes (Fig. 1). Whether measured by optical density or by use of a ribosomal RNA-specific cloned probe, recruitment starts immediately after fertilization and proceeds in a linear manner for ~90 min before reaching a plateau at 20–25% of the available rRNA. In

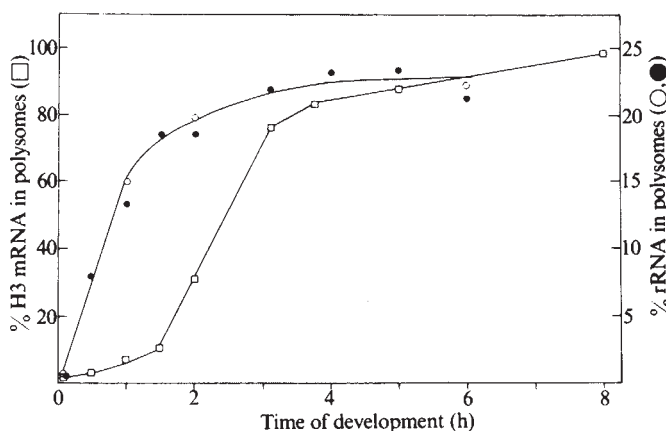


Fig. 1 Time of appearance of rRNA and histone H3 mRNA in the polysomes of *S. purpuratus* embryos. Embryos were cultured at 13.5 °C for the appropriate lengths of time and polysomes were prepared by the procedure of Wold *et al.*¹⁰, omitting the polyvinyl sulphate. Material was layered on to 5–40% exponential sucrose gradients and centrifuged for 2.5 h, 26,000 r.p.m. in an SW27 rotor. Individual fractions were collected, deproteinized with phenol/chloroform/isoamyl alcohol (25:24:1) and RNA precipitated with ethanol. RNA samples were taken up in 1×MOPS buffer (20 mM morpholinopropane sulphonic acid, 5 mM sodium acetate, 1 mM EDTA) at pH 7.0, 6% formaldehyde and 50% formamide, heated at 65 °C for 5 min, cooled rapidly and electrophoresed on a 1% agarose gel in 1×MOPS buffer containing 6% formaldehyde. After electrophoresis the gel was soaked for 30 min in 20×SSC (1×SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.4), and blotted on to a nitrocellulose filter in 10×SSC. The filter was rinsed with 3×SSC and baked for 2 h at 80 °C in a vacuum oven. Probes for histone H3 (p3H1, provided by L. Kedes) and rRNA (pLv 60, provided by D. Stafford) were nick-translated to 1–5 × 10⁷ c.p.m. μg⁻¹ using [α -³²P]dCTP. Filters were prewashed in 10×Denhardt's solution (0.02% bovine serum albumin, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone), 0.1% SDS, 0.02 M phosphate buffer, 5×SET (1×SET = 0.15 M NaCl, 0.025 M Tris, pH 8.0, 0.002 M EDTA) at 68 °C for 1–2 h. The filters were prehybridized at 45 °C in 10 ml of 10% dextran sulphate, 5×SET, 0.02 M phosphate buffer, 1×Denhardt's, 100 μg ml⁻¹ sheared calf thymus DNA, 100 μg ml⁻¹ poly (rA) and 50% formamide. After 1–2 h, 1–5 × 10⁷ c.p.m. of nick-translated plasmid was added and hybridized at 45 °C for 12–16 h. After hybridization, the filters were washed at 68 °C for 1 h in each of the following: (1) 5×SET, 0.1% SDS, 0.1% sodium pyrophosphate and 0.025 M phosphate buffer; (2) 1×SET, 1×Denhardt's, 0.1% sodium pyrophosphate and 0.025 M phosphate buffer; (3) 0.3×SET, 0.025 M phosphate buffer, 0.1% sodium pyrophosphate, 0.1% SDS. The filters were air dried and exposed to preflashed¹¹ Kodak X-Omat RP film at -70 °C using Dupont Cronex lighting plus intensifier screens. Exposure times ranged from 2 to 96 h. ●, % rRNA in polysomes as determined by A₂₅₅; □, % rRNA in polysomes as determined by hybridization with rDNA probe pLv60; □, % of histone H3 mRNA in polysomes.

addition, titrations with labelled poly(U) show that >70% of the total cytoplasmic poly(A) is present in the polysomes of 2-h-old embryos, while <5% is found in polysomes of unfertilized eggs⁶.

To determine whether H3 mRNA is recruited with the same kinetics as total mRNA, we have measured the relative amount of H3 mRNA present in the mRNA and polysome compartments of the embryo from fertilization through 8 h of development. *Strongylocentrotus purpuratus* embryos were cultured for specific times and the polysomes displayed on 5–40% exponential sucrose gradients. After fractionation the RNA was extracted, electrophoresed on agarose gels and blotted on to nitrocellulose filters. The filter-bound RNA was then hybridized to a ³²P-labelled nick-translated clone containing only the coding sequence for histone H3. The amount of probe hybridizing to each region of the gradient gave a direct measure of the percent of H3 mRNA found in the polysomal and mRNA compartments of the embryo at a given time (Fig. 2a, c). Quantification was done by densitometric scanning of autoradiographs exposed within the linear range of the film. To ensure that the detected RNA was polysomal, parallel experiments were performed using EDTA to dissociate polysomes (Fig. 2b, d).

At 90 min after fertilization, <10% of the total hybridizable H3 mRNA is found in the polysomal region of the gradient (Fig. 1). Between 90 min and 4 h there is a rapid linear rise in the percent of the total H3 mRNA in the polysomes until, after 4 h, >85% of the H3 mRNA is detected in the polysome region of the gradient. Whether embryos are cultured in the presence or absence of actinomycin D, we can detect no significant increase in the amount (per μ g of total RNA) of H3 mRNA during the first 5 h after fertilization (Fig. 3). This, in conjunction with the recruitment data in Fig. 1, suggests that little of the H3 mRNA initially present in the egg remains in the masked state. Newly synthesized H3 mRNA begins to accumulate measurably 6–8 h after fertilization, and seems to move rapidly into the polysomal fraction, as by 8 h >95% of the H3 mRNA is present in polysomes. In contrast to H3 mRNA, α -tubulin mRNA does not exhibit the 90-min delay, but is significantly loaded by 30 min after fertilization (data not shown).

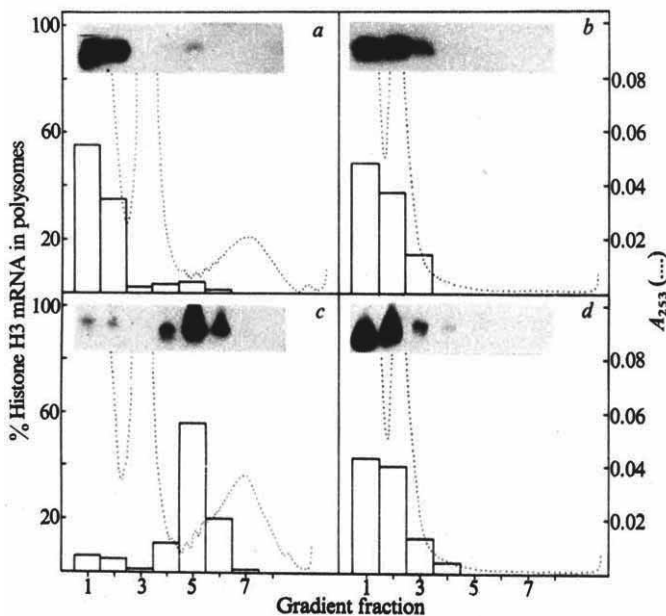


Fig. 2 Autoradiographs and quantification of representative H3 mRNA titration from 1-h and 3.75-h embryos. The amount of probe hybridizing to each fraction of the gradient is expressed as the percentage of the sum of the probe hybridized in all fractions. A_{253} (.....) was obtained using an Isco UA5 absorbance monitor. Fractions were collected from the top. a, 1-h embryos; b, 1-h embryos, EDTA released; c, 3.75-h embryos; d, 3.75-h embryos, EDTA released.

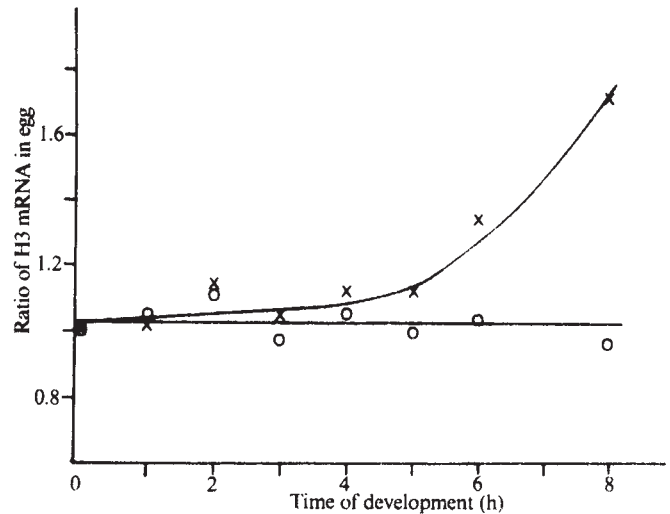


Fig. 3 Titration of total H3 mRNA from unfertilized egg through 8 h of development in the presence or absence of actinomycin D. Embryos were cultured in the presence of $25 \mu\text{g ml}^{-1}$ actinomycin D. At each time point, identical aliquots were removed and the total RNA extracted, electrophoresed, blotted and hybridized as described in Fig. 1 legend. Data are given as the ratio of the amount of H3 at each time point to the amount in the unfertilized egg. x, Control; o, actinomycin D.

Consistent with these results is a recent study of *in vitro* translation of sea urchin mRNA using a cell-free system derived from sea urchin eggs. This study shows that histone H3, as well as other histones, are synthesized *in vitro* in significant amounts; however, very little histone synthesis was observed *in vivo* until at least 1 h after fertilization⁷. Our data show that this delay in appearance of histone H3 is to be expected from the behaviour of the H3 mRNA. Because most mRNAs in the egg must be unmasked before translation can occur⁸, the delay in H3 mRNA recruitment may be due to a specific delay in unmasking of histone mRNAs. Alternatively, the delay may be due to a change in initiation properties of ribosomes or the lack of a histone mRNA-specific translation factor. We cannot yet discriminate between these possibilities. Woods and Fitschen⁹ have observed changes in histone mRNA distribution in sea urchin polysomes during telophase of the first cleavage cycle and suggest that it may be associated with the onset of DNA synthesis. However, the absence of any further association of use of histone mRNA with the cell cycle argues against a direct causal association between DNA synthesis and histone mRNA recruitment. The marked contrast between the recruitment of H3 mRNA and that of most maternal mRNAs clearly demonstrates translational regulation of early embryonic protein synthesis in the sea urchin embryo.

This work was supported by USPHS grants HD06902 to R.A.R. and HD14182 to W.H.K. D.E.W. is a pre-doctoral trainee supported by USPHS award T32 GM07227. We thank L. H. Kedes, D. Stafford and D. Mischke and M. L. Pardue for gifts of clones used in this work.

Received 8 April; accepted 27 May 1981.

1. Raff, R. A. in *Cell Biology: A Comprehensive Treatise* Vol. 5, 107–136 (Academic, 1980).
2. Brandhorst, B. *Dev. Biol.* **52**, 310–317 (1976).
3. Rosenthal, E. T., Hunt, T. & Ruderman, J. V. *Cell* **20**, 487–494 (1980).
4. Davidson, E. H. *Gene Activity in Early Development*, 2nd edn (Academic, New York, 1976).
5. Raff, R. A., Brandis, J. W., Huffman, C. J., Koch, A. L. & Leister, D. E. *Dev. Biol.* (in the press).
6. Dolecki, G. J., Duncan, R. F. & Humphreys, T. *Cell* **11**, 339–344 (1977).
7. Winkler, M. & Steinhardt, R. *Dev. Biol.* (in the press).
8. Jenkins, N. A., Kaumeyer, J. F., Young, E. M. & Raff, R. A. *Dev. Biol.* **63**, 279–298 (1978).
9. Woods, D. E. & Fitschen, W. *Cell Diff.* **7**, 103–114 (1978).
10. Wold, B. J., Klein, W. H., Hough-Evans, B. R., Britten, R. J. & Davidson, E. H. *Cell* **14**, 941–950 (1978).
11. Laskey, R. A. & Mills, A. D. *Eur. J. Biochem.* **56**, 335–341 (1975).

BOOK REVIEWS

In pursuit of a self-conscious science

J. O'Keefe and P.D. Wall

FOR the half-century between 1920 and 1970, most neurophysiologists and experimental psychologists were content to leave "consciousness" outside the boundaries of their disciplines. The properties traditionally attributed to consciousness — freedom of the will, creativity, self-knowledge — seemed difficult, if not impossible, to capture within the accepted notions of deterministic science. Furthermore there was a strong reaction to the use of consciousness as an explanatory device: two actions were different because one was voluntary; two stimuli differed because one was attended to.

In recent years this radical behaviourist approach has been modified and recent work has thrust the problem of consciousness to the forefront again. One scientist who deserves considerable credit for insisting on a place for consciousness in neurophysiology during this scientific dark night of the soul is the author of the book under review. In 1953, Sir John Eccles published a book entitled *The Neurophysiological Basis of Mind* (Oxford University Press) which was quite the best summary up to that time of how nerve impulses pass from one cell to another. Having been for many years one of the main proponents of the electrical synapse theory, Eccles accepted the new evidence for the chemical synapse and became one of the architects of our modern understanding of neuronal transmission, for which work he received the Nobel prize in 1964.

In 1953 virtually no aspect of brain research had a direct bearing on the question of consciousness and Professor Eccles's inclusion of a final chapter on this topic was regarded as a mild eccentricity. Consciousness entered the scientific picture as a subtle force operating at the level of the indeterminate synapse. Eccles's preoccupation with the brain-mind problem has increased over the years: the last chapter of the 1953 book has become the first chapter of the present book. The book is divided into three sections. The first section (Lectures 1 and 2) sets out Eccles's current theory of the mind which he has developed in association with the philosopher Sir Karl Popper. The second (Lectures 3–6) describes recent experiments in neurophysiology and psychology which bear on the question of consciousness, and the third (Lectures 7–10) marks a shift to a consideration of

The Human Psyche. The Gifford Lectures, University of Edinburgh 1978–1979. By John C. Eccles. Pp.279. ISBN 3-540-09954-9. (Springer-Verlag: 1980.) DM44, \$26.

the broader issues of the relation between consciousness and social values.

Eccles's theory of consciousness views it as an entity which exists independently of the brain but which interacts strongly with it. He rejects materialist theories which identify consciousness with the activity of a part or all of the brain. The brain-consciousness interaction takes place at the level of the cerebral cortex which unfortunately does not have the appropriate anatomical structure. Whereas consciousness appears to be a unity which maintains its integrity and continuity over time, the cerebral cortex consists of numerous, relatively independent modules each coding for a different aspect of the sensory environment or motor behaviour. Eccles concludes that "... the unity of conscious experience is provided by the self-conscious mind and not by the neural machinery of the liaison areas of the cerebral hemisphere" (p.50). How is this accomplished?

... the self-conscious mind can scan the activity of each module of the liaison brain or at least those modules tuned in to its present interest. ... the self-conscious mind has the function of integrating its selections from the immense patterned input it receives from the liaison brain. ... in order to build its experiences from moment-to-moment [p.46].

It would appear that on this formulation the problem of consciousness is inaccessible to neurophysiological analysis and perhaps to any scientific analysis. It is a non-neural entity with its own structure, interests, expectations and creative impulses, attributes which are never discussed but which are constantly invoked to "explain" psychological and physiological findings.

Much of the middle section of the book is an excellent description of recent findings. As might be expected from the author's preoccupation with the cerebral cortex as the part of the brain which liaises with the self-conscious mind, much (and by far the most satisfactory part) of this section deals with effects ascribable to the cortex. Visual illusions show that much of the sensory world is created by the brain

from inadequate data; EEG recordings of the brain potentials from the scalp show that cortical activity precedes voluntary movements and accompanies anticipated stimuli. Unfortunately, these fascinating findings are "explained" by hand-waving references to the self-conscious brain. Since the "explanation" is given in terms of a world with normal scientific rules, one may reasonably ask for experiments to investigate this world, but as we approach it retreats as a will'o the wisp.

Subsequent chapters give a cursory but adequate account of the limbic system, the pain system and the reticular activating system of the brainstem. Professor Eccles acknowledges that these may have some bearing on the problem of consciousness but keeps his sights firmly fixed on the cerebral cortex. A surmise in the section on the limbic mechanisms of emotion "that Gary Gilmour, the last criminal to be executed in the United States, may have been suffering from amygdaloid seizures" (p.116–117) warns us of what to expect in the third and final section of the book: the relation between brain, consciousness and social values.

This last section contains a wide ranging set of comments and quotations on topics centred around the twilight area where scientific ideas impinge upon social values. Professor Eccles is particularly severe on scientists who would reduce human action or human values to mechanistic explanations. The sociobiologist E.O. Wilson who seeks to explain altruistic behaviour on the basis of selfish genes and the molecular biologist Jacques Monod who seeks to emphasize the chance elements in human existence come in for special opprobrium. Although some cogent arguments are advanced, much of the discussion takes the form of poetic quotations from suitably authoritative sources: Sherrington, Planck, Heisenberg, Eddington and Einstein. We suggest that most readers will take these last chapters for what they are: the personal prejudices of a great neurobiologist who has strayed outside the bounds of his area of expertise. We hope it won't deter them from giving serious consideration to the earlier chapters.

J. O'Keefe and P.D. Wall are members of the Cerebral Functions Research Group at University College, University of London.

A tool of elegance and great power for physicists . . .

Peter McClintock

Green's Functions and Condensed Matter. By G. Rickayzen. Pp.357. ISBN 0-12-587950-4. (Academic: 1980.) £22.80, \$55. *Many-Particle Physics.* By G.D. Mahan. Pp.1,003. ISBN 0-306-40411-7. (Plenum: 1981.) \$85, £53.55.

WHEN George Green, miller and part-time theoretical physicist, died in 1841 at the early age of 47 the *Nottingham Review* commented regretfully that "... had his life been prolonged, he might have stood eminently high as a mathematician". Little did they realize. During his brief but remarkably productive twelve-year spell of scientific activity, starting with *An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism* (published privately in 1828), Green had introduced the concept of electric potential, had carried out work of fundamental importance in wave theory and hydrodynamics and had laid the foundations for modern theories of elasticity. Notwithstanding this substantial contribution, however, and despite his acknowledged influence on both Stokes and Kelvin, Green's memory remained relatively obscure until more than a century after his death, when R.P. Feynman took up some of the techniques introduced by Green, and developed them for application to nuclear physics and the theory of elementary particles. The use of Green's function methods has subsequently spread through the rest of physics including, particularly, solid state physics.

Thus it has come about that no physicist in the 1980s can remain indifferent to Green's functions. For those who love them including, probably, the majority of theorists, Green's functions constitute a tool of elegance and great power, indispensable in some areas of physics, which can with advantage be applied to almost any problem with which they may be confronted. Others regard them with profound suspicion. They feel that devotees frequently employ Green's function techniques for their own sake, for the sheer joy of using them, and often at the expense of physical insights which can more readily be gained through simpler approaches. Their mistrust probably relates, at least in part, to a lack of understanding. The publication of two new advanced textbooks on the subject, each aimed at filling what has been an unfortunate lacuna in the literature, is therefore to be applauded.

The similarities between G. Rickayzen's *Green's Functions and Condensed Matter* and G.D. Mahan's *Many-Particle Physics* are greater than their differences. Both books are aimed at advanced graduate students, or academic staff, and both assume a good prior knowledge of quantum mechanics at the level provided,

say, by the classic texts of Schiff or Landau and Lifshitz. They are both most definitely books of physics, as opposed to mathematics, each consisting very largely of an account of how Green's function and Feynman diagram techniques can be applied in practice to a selection of topics in solid state physics. Both books cover, for example, phonons, the Coulomb gas, electron-phonon interactions, transport theory, linear response, superconductivity and liquid helium (which by tradition is treated as solid state physics). Both provide, for each chapter, references to the original literature (Mahan's being the more extensive) and a set of problems, without solutions. In each case, the standard of production is high. The most obvious difference between the two books lies in their respective lengths: Rickayzen's has 357 pages in total whereas Mahan's, with 1,003, is almost three times as long. Much of this difference can be accounted for in terms of Mahan's detailed treatment of certain topics — for example, optical properties or polarons — which Rickayzen either omits or treats less fully; but, to a considerable extent, it also relates to their very different styles of presentation.

Mahan's stated objective was

... to take standard subjects ... and to summarize what is generally known. All the steps are retained in the derivation, so that the answers are obtained by starting from the beginning and working through to the end.

This latter feature is quite unusual in advanced texts, which commonly leave the reader to fill in a lot of the intermediate steps for himself, and it is a reflection of Mahan's declared intention of producing a book which graduate students will be able to use on their own, if necessary, in the

absence of a formally taught course. It will also be a considerable help and encouragement to the less theoretically biased readers at all ages and stages. Mahan writes in a style which is relaxed without being imprecise, and which often displays refreshing candour as, for example, in relation to the ground state properties of quantum fluids where he comments that, unusually, "... the Green's function method gives awful results".

Rickayzen's book, which starts disarmingly with the truism that "Anyone who has used the Coulomb potential due to a point charge has used a Green's function", is by contrast a model of brevity and succinctness. The most immediate benefit lies in the remarkable range of topics which he has managed to treat at a useful level in the space available: in addition to the items mentioned above, the book also includes chapters on magnetism, disordered systems and critical behaviour. The relatively condensed presentation should be well suited to the graduate student in theoretical physics at whom the book is primarily aimed, particularly if used to support and supplement the material provided in a formal course of lectures on the subject.

Each of these books is excellent in its own way. They should be a real help both in educating the next generation of solid state theorists and also in rendering "Greenery" just a little bit less daunting to those professional physicists who have not yet fully come to terms with the power and utility of these techniques. □

Peter McClintock is a Senior Lecturer in the Department of Physics, University of Lancaster.

... and another for molecular biologists

Ueli Schibler

Genetic Engineering 1. Edited by Robert Williamson. Pp.168. ISBN 0-12-270301-4. (Academic: 1981.) £9.80, \$24.

GENETIC engineering has not turned dull molecular biologists into imaginative ones. There is, however, little doubt that both could profit immensely from this elegant and powerful technique. This is enough justification for a new series entitled *Genetic Engineering*. According to the editor, the purpose of the new periodical is to rapidly publish comprehensive reviews about different aspects of recombinant DNA technology. This first issue addresses a very heterogeneous readership, including students, experienced researchers and physicians. It brings together a rather technical article about cDNA cloning, an

article about prenatal diagnosis of abnormal haemoglobin genes and a review on transcription of cloned genes in different experimental systems.

J. E. Williams provides a meticulous account of every aspect of the applicability and feasibility of cDNA cloning, and evaluates most of the tricks used in the enzymatic construction of recombinant molecules containing cDNA and their propagation in bacterial host cells. Particularly helpful are the discussions of mRNA complexity and the size of a cDNA library. Using this information one can easily calculate how many bacterial colonies should be screened to obtain, with a certain probability, an mRNA sequence of a given abundance. It is a pity that Williams did not include a short protocol

which he has found to be generally applicable. This might have reduced the beginner's confusion over choosing the most appropriate procedures, since he has just read 55 pages describing alternative methods, loop-holes and pitfalls.

The second article deals with a medical application of recombinant DNA technology. P. F. R. Little summarizes the lesions of haemoglobin genes associated with different haemoglobin deficiencies and how they can be detected by Southern blotting of DNA from fetal cells obtained by amniocentesis. Two major principles are discussed: the first is direct demonstration of deletions within a globin gene itself; this approach is particularly valuable for certain thalassaemias. The second is based on the fact that deficient genes are often linked to polymorphism of nearby restriction endonuclease sites, which in turn is readily detectable. This approach can be applied to any dysfunctional gene where linkage exists, and does not depend on the nature of the lesion in the DNA. As, thanks to molecular cloning technology, our knowledge about genes increases rapidly, the antenatal diagnosis of lesions at the DNA level will certainly have an important future in preventive medicine.

Several systems to study the expression of cloned genes either *in vitro* or *in vivo* have recently become available. Two of them, the transcription of isolated genes in microinjected oocytes, and in cell-free extracts, are reviewed in an authoritative article by M. P. Wickens and R. A. Lasky. Most of the important published work carried out on the expression of eucaryotic genes in these systems is critically reviewed and put into perspective. *In vitro* transcription of genes, both wild type and after mutation *in vitro*, has been important in recognizing signals on DNA. This strategy succeeded in defining promoters for various genes transcribed by RNA polymerase III, and sequence elements in genes transcribed by RNA polymerase II which are required for correct initiation and modulation of transcription.

All three articles of the first issue of *Genetic Engineering* are well written and certainly worth reading. It seems to me that the articles which are published in this issue and are planned for publication in further issues could have been grouped more logically, for example by putting the construction of cDNA (this issue) and genomic DNA recombinants (planned for a future issue) into the same volume. The same holds true for Wickens and Lasky's article and the planned article on expression of cloned genes in transfected cells. However, if the promised publishing strategy is adhered to, the series should quickly build into a useful reference work on recombinant DNA technology. □

Ueli Schibler is a Junior Staff Member at the Swiss Institute for Experimental Cancer Research, Lausanne.

Isolation and assay of insect hormones

H. F. Nijhout

Neurohormonal Techniques in Insects. Edited by Thomas A. Miller. Pp.282. ISBN 3-540-90451-4. (Springer-Verlag: 1981.) DM79, \$46.70.

IF IT is possible to pinpoint the birth of a biological discipline, then neuroendocrinology was born in 1917 when a Polish biologist, Stefan Kopeč, reported that the brain of gypsy moth caterpillars produced a hormone that stimulated pupation. The hormone he discovered we now call prothoracicotropic hormone (PTTH), because it stimulates the prothoracic glands to secrete the moulting hormone, ecdysone. In the decades that have passed since this landmark discovery, the recognition, isolation and identification of additional insect neurohormones has proceeded at a slow but steady pace and *Neurohormonal Techniques in Insects* provides an account of the current status of these endeavours.

The purpose of this volume is to provide a compendium of the methods that are currently used by various laboratories to collect, assay and purify the neuroendocrines of insects. The book is divided into 11 chapters dealing with as many neuroendocrine factors, to wit: proctolin; adipokinetic hormone; diuretic hormone; insulin- and glucagon-like hormones; bursicon; pupariation factors; cuticle plasticizing factors; eclosion hormone; diapause hormone and PTTH. Each chapter has a brief introduction establishing the biological properties of the hormone in question, followed by a review of various bioassay methods and finally a detailed description of the methods used in its isolation and characterization.

The first two chapters by Staratt and Steele and by Stone and Mordue on the isolation and identification of proctolin and adipokinetic hormone, respectively, are by far the best methods papers I have ever seen. The quality of the remaining contributions is, predictably, variable both in method of exposition as well as in the skill with which various problems are handled. Some of the defects become understandable, though, when one considers the enormous obstacles that must be overcome by those who discover a new hormone and venture to identify it. The numbers of insects that need to be worked up to extract a sufficient quantity of hormone are truly prodigious. For example, 125,000 cockroaches yielded 1.12 mg of proctolin, and fractionation of 96,000 silkworm heads yielded 3 mg of a heterogeneous residue possessing high PTTH activity. To make matters worse, the bioassays that must be used are cumbersome and consume vast quantities of time, manpower and hard-won hormone. The biggest problem, though, is that the specificity of a bioassay for a given

hormone is often questionable, and I found a critical consideration of what constitutes an optimal assay method and scoring procedure lacking in most chapters.

A budding biochemist should have no difficulty in following the rationale as well as the procedures employed by the various authors and it should be possible for any well equipped laboratory to employ the methods described to attempt their own hormone isolations. However, it is evident from the accounts in this volume that the development of reliable, and preferably biochemical, assays should be the first order of business for the future. □

H. F. Nijhout is Associate Professor of Zoology at Duke University, Durham, North Carolina.

Creative catalysis

G.W. Parshall

Homogeneous Transition-metal Catalysis: A Gentle Art. By Christopher Masters. ISBN hbk 0-412-22110-1; ISBN pbk 0-412-22120-9. (Chapman & Hall: 1980.) Hbk £20, \$25; pbk £9.50.

THE gentle art described in this book is the synthesis of organic compounds through the skilful application of transition metal complexes as catalysts. Dr Masters' purpose in writing this book was to stimulate enthusiasm for this art. He succeeds admirably because he has produced not just a readable description of the subject as it exists today but also a perceptive analysis of developing research areas.

In the introductory chapter, Dr Masters describes the basic principles of bonding and reactivity in transition metal complexes in simple terms for the non-specialist. Phenomena such as oxidative addition, β -elimination and steric and electronic effects are illustrated with significant examples. The fundamentally kinetic nature of catalysis is stressed.

The heart of the book is a description of homogeneous catalysis as it is practised today. Fourteen of the two dozen or more major commercial applications of soluble transition metal catalysts are discussed in some detail. The processes described include the reactions of carbon monoxide with acetylenes, olefins and alcohols, the dimerizations of olefins and dienes, and the oxidations of several different types of hydrocarbons. Fundamentally organometallic processes such as the hydrogenation, isomerization, polymerization and metathesis of olefins are also included, even though industry usually chooses heterogeneous catalysts for these

operations. Some significant reactions like the hydrocyanation and hydrosilylation of olefins are omitted and others are simply mentioned. However, the reactions that are included are discussed expertly and with sufficient detail to satisfy the reader who is just becoming acquainted with the field.

The book closes with analysis of research trends in homogeneous catalysis. Nitrogen fixation, C-H bond activation and CO hydrogenation are cited as well-studied areas in which the crucial discoveries have yet to be made. Other topics such as bimetallic catalysis are suggested as profitable areas for research.

Homogeneous Transition-metal Catalysis will probably be used largely as a reference or introductory text for the

industrial chemist working in this field. It may also be used in teaching graduate-level courses in catalysis or organometallic chemistry. The emphasis on understanding and on research should appeal to the student.

Dr Masters has succeeded nicely in his objective to convey enthusiasm for homogeneous catalysis. It is a dynamic subject from both scientific and practical viewpoints and this emerges clearly from the text. The reader should be both stimulated and informed by perusal of this book. □

G. W. Parshall is Director of Chemical Science in the Central Research and Development Department of E. I. du Pont de Nemours & Co., Wilmington, Delaware. His most recent book is Homogeneous Catalysis (Wiley, 1980).

colleagues and consultants. Written by a single author, the book's structure and style are unified and integrated to a degree that multi-author publications rarely achieve.

Brown divides his subject matter into 12 chapters, two of which form essentially the first half of the book and provide a systematic synopsis of African freshwater snails; two short chapters cover snail-borne parasites of medical and veterinary importance in Africa, and the remainder deal with ecological, physiological and biogeographical aspects of snail biology, stressing snail hosts of schistosomes and approaches to snail population control.

The synopsis of more than 300 gastropod species is efficiently and accurately presented, including carefully documented descriptions of morphological and ecological characteristics, plus (usefully) a brief statement of their parasitological importance as real or potential intermediate hosts. The text is thoroughly illustrated with 128 photographic plates, line drawings and distribution maps, all arranged to maximum advantage. Here, as in all chapters, a comprehensive, well-selected citation of literature stresses work done in Africa or on material of African origin; inclusion of a large number of contributions from publications in southern Africa will be of special interest to readers unfamiliar with the volume and quality of this body of literature.

The two chapters on snail-borne parasites of medical and veterinary importance concentrate on snail-parasite interrelationships. The account of snails and schistosomes is, justifiably, given prominence since the impact of schistosomiasis on human beings is so great, and because so much more data is available for discussion of this complex. One deficiency in this section is the omission of the human disease, heterophyiasis (admittedly referred to in the systematic description of its snail host, *Pirenella*), in favour of veterinary parasites such as paramphistomes, and a potentially occurring human disease, angiostrongylosis. Overall, I felt that the parasitological section was disappointingly brief, seeming as it did to be included as a kind of (unnecessary) justification for the rest of the book.

The remaining six chapters on ecological, physiological and biogeographical aspects of African snails are excellent; particularly outstanding is that on the biology of the genus *Bulinus*, a subject of great complexity and interest to biomedical scientists. Completing the book are a unique appendix, in which major methods of field and laboratory studies of snails and parasites are presented, and taxonomic and general subject indexes; I found the latter extremely useful and accurate, and reflective of the general editorial excellence of the book.

Frank J. Etges is Professor of Parasitology at the University of Cincinnati.

Illumination on the light reactions

J. Barber

Photosynthesis: Physical Mechanisms and Chemical Patterns. By Roderick K. Clayton. Pp.275. ISBN hbk 0-521-22300-8; pbk ISBN 0-521-29443-6. (Cambridge University Press: 1981.) Hbk £17.50, \$32.50; pbk £6.95, \$11.95.

CONSISTENT with Roderick Clayton's outstanding talents as a scientist and author, I judge this, his latest book, as being first class. He has written a lucid and well-structured text which makes marvellous reading for the expert and also will satisfy the requirements of non-specialists and students. He achieves this, in part, by "digressions" and by giving background "notes" in an appendix. The overall effect is to pass on the excitement of an active scientist who, with the qualities of a talented teacher, instinctively takes the reader through logical developments calling on the basic laws of photophysics, photochemistry and thermodynamics wherever necessary.

From the title it is clear that Clayton has concentrated on the "light reactions" of photosynthesis and consequently has restricted his discussion of carbon fixation (the "dark reactions") to one short chapter at the end of the book. It is inevitable that in producing a modern text on the light reactions of photosynthesis he should not only deal with oxygen-evolving organisms, but also with photosynthetic bacteria. The book has been divided into four distinct parts, each consisting of several chapters. Part I gives a historical account which introduces the reader to the basic concepts of the light reactions, including the nature and properties of the pigments and the existence of different types of photosystems. Part II then concentrates on the organizational aspects, while Part III deals with the photochemistry and associated transfer of electrons and protons. Part IV

is concerned mainly with photophosphorylation. At the end of each chapter the author has listed a few selected references for further reading. Also, the background notes in the appendix cite other key work.

Without any hesitation, I thoroughly recommend this book to specialists and non-specialists alike who have an interest in the more biophysical aspects of photosynthesis, whether it be from the research or the teaching point of view. I also think that those working in other areas of photobiology may find much of the basic substance of the book useful. □

J. Barber is Professor of Plant Physiology in the Department of Pure and Applied Biology, Imperial College, University of London.

Medical malacology

Frank J. Etges

Freshwater Snails of Africa and their Medical Importance. By David S. Brown. Pp.487. ISBN 0-85066-145-5. (Taylor & Francis, London/American Malacologists Inc., Melbourne, Florida: 1980.) £25, \$55.

DAVID BROWN is to be congratulated on writing an exceptionally fine treatise on the gastropod fauna of Africa, a reference book which will be set a standard of quality and content for many years. His stated goal of presenting a "comprehensive" account of this subject which will be of value to specialists and students of taxonomy, medical malacology and freshwater biology has surely been attained. This achievement is a tribute both to the broad experience of the author, and to the abilities of his many acknowledged